

## Research Article

# N-terminal acetylation in a third protein family of vertebrate alcohol dehydrogenase/retinal reductase found through a ‘proteomics’ approach in enzyme characterization

D. Hirschberg<sup>a</sup>, E. Cederlund<sup>a</sup>, B. Crosas<sup>b</sup>, A. Jonsson<sup>a</sup>, S. Tryggvason<sup>a</sup>, J. Farrés<sup>b</sup>, X. Parés<sup>b</sup>, T. Bergman<sup>a</sup> and H. Jörnvall<sup>a,\*</sup>

<sup>a</sup> Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm (Sweden), Fax +46 8 337 462, e-mail: hans.jorvall@mbk.ki.se

<sup>b</sup> Department of Biochemistry and Molecular Biology, Universitat Autònoma de Barcelona, 08193 Bellaterra (Spain)

Received 27 June 2001; accepted 2 July 2001

**Abstract.** A recent finding of a novel class of retinol-active alcohol dehydrogenase (ADH) in frog prompted analysis of this activity in other vertebrate forms. Surprisingly, yet another and still more unrelated ADH was identified in chicken tissues. It was found to be a member of the aldo-keto reductase (AKR) enzyme family, not previously known as an ADH in vertebrates. Its terminal blocking group and the N-terminal segment, not assigned

by protein and cDNA structure analysis, were determined by electrospray tandem mass spectrometry after protein isolation by two-dimensional gel electrophoresis. The N terminus is Acetyl-Ala- and the N-terminal segment contains two consecutive Asn residues. The results establish the new ADH enzyme of the AKR family and show the usefulness of combined gel separation and mass spectrometry in enzyme characterization.

**Key words.** Two-dimensional gel electrophoresis; proteomics; tandem mass spectrometry; N-terminal acetylation; Aldo-keto reductase; enzyme isolation.

Elucidation of structure-function relationships and catalytic mechanisms of novel enzymes generally involves successive steps of protein purification, enzyme characterization, partial sequence analysis, and molecular cloning. In cases with novel mechanisms, multiple forms, or special contaminants, protein isolation may be difficult and analysis may start without pure material. One such case was encountered during a recent definition of an aldo-keto reductase (AKR) as an alcohol dehydrogenase (ADH) with retinal reductase activity in chicken tissues [1]. Usually, vertebrate ADHs are of the medium-chain dehydrogenase/reductase (MDR) zinc-containing protein type [2, 3]. Alternatively, several steroid dehydro-

genases and other ADHs are of the short-chain dehydrogenase/reductase (SDR) family [2, 4]. Enzymes of both families are active with retinoids, and are believed to play a key role in regulating the production of retinoic acid [5, 6]. However, the initial sequence data suggested that the novel ADH was of a third family type, AKR [1], previously not known to represent ADH or retinal reductase activities, but rather aldehyde and aldose reductases [7]. This chicken protein was characterized by a combination of protein and cDNA sequence analysis, correlated with model building and enzymatic analysis, which yielded information on all segments of the primary structure except the N-terminal part [1].

To verify the new structure, and to establish the nature of the blocked, chemically modified N terminus, two-dimensional (2-D) gel electrophoresis combined with tan-

\* Corresponding author.

dem mass spectrometry (MS) was utilized, demonstrating the usefulness of a 'proteomics' approach [8] also in conventional enzyme characterization.

## Materials and methods

### Protein material and 2-D gel separation

The novel ADH was prefractionated from chicken upper digestive tract tissues by column chromatography steps and carboxymethylated as described [1]. For final purification, 2-D gel electrophoresis was employed. The protein solution (20  $\mu$ l containing 0.04  $\mu$ g/ $\mu$ l) from the last chromatography step was dried under a stream of nitrogen and subsequently dissolved in 150  $\mu$ l lysis buffer containing 8 M urea, 4% (w/v) CHAPS, 1% (w/v) DTT, 0.6% (v/v) Triton X-100, and 2% (v/v) immobilized pH gradient buffer (pH 3–10, nonlinear; Amersham Pharmacia Biotech). Isoelectric focusing was carried out on a 7-cm, pH 3–10 nonlinear Immobiline Drystrip (Amersham Pharmacia Biotech) using 50  $\mu$ l of the sample solution. The voltage was increased from 0 to 500 V for 2 h, after which it was further increased to 3500 V for an additional 1.5 h, and then kept constant at 3500 V for 4.5 h. The second-dimension SDS-PAGE was performed in a 10% polyacrylamide gel and was stained with Coomassie brilliant blue (R250).

### In-gel digestion and mass spectrometry

Protein spots were excised and washed at 37 °C in 200  $\mu$ l 100 mM ammonium bicarbonate containing 50% (v/v) acetonitrile, until the Coomassie stain was removed. The supernatant was discarded and the gel pieces were shrunk using 100  $\mu$ l acetonitrile twice and subsequently dried for 20 min in a vacuum centrifuge. Trypsin (modified trypsin; Promega) was dissolved in 200 mM ammonium bicarbonate to a final concentration of 0.1  $\mu$ g/ $\mu$ l. After incubation with 5  $\mu$ l trypsin solution for 10 min at room temperature, 50  $\mu$ l 200 mM ammonium bicarbonate was added and the pieces were further incubated overnight at 37 °C. Digestion was stopped with 1  $\mu$ l trifluoroacetic acid (TFA). The supernatant was collected. The peptides were extracted from the gel matrix twice with 100  $\mu$ l 60% acetonitrile/0.1% TFA for 1 h at 37 °C, followed by one extraction with 50  $\mu$ l 40% acetonitrile/0.1% TFA for 10 min at 37 °C, and finally with 20–30  $\mu$ l acetonitrile at room temperature. The supernatant and the extracts were pooled and concentrated to 10–20  $\mu$ l under a stream of nitrogen. The combined peptide extract was desalted using C<sub>18</sub> ZipTip (Millipore). The ZipTip was activated and equilibrated using 10  $\mu$ l 70% acetonitrile/0.1% TFA, twice, 10  $\mu$ l 50% acetonitrile/0.1% TFA, twice, and 10  $\mu$ l 0.1% TFA twice. The sample was loaded onto the ZipTip by pipetting the peptide extract 20 times, and subsequently washed using 10  $\mu$ l 0.1% TFA twice. The tryptic

fragments were then eluted using 60% acetonitrile/1% acetic acid for subsequent analysis by tandem MS, or 75% acetonitrile/0.1% TFA for subsequent matrix-assisted laser desorption/ionization (MALDI) MS.

The tryptic fragments were analyzed by MALDI MS (Voyager DE-PRO; Applied Biosystems) using  $\alpha$ -cyano-4-hydroxycinnamic acid as matrix (saturated solution in 70% acetonitrile) by mixing 1:1 (v/v) with the sample. Amino acid sequences were determined using electrospray ionization (ESI) in a quadrupole time-of-flight (Q-TOF) tandem mass spectrometer (Micromass), equipped with an orthogonal sampling ES-interface (Z-spray; Micromass). Samples were introduced via gold-coated nano-ES needles (Protana). A capillary voltage of 800–1000 V was applied together with a cone voltage of 40–45 V and a collision energy of 4.2 eV. The sample aerosol was desolvated in a stream of nitrogen. During collision-induced dissociation of the tryptic fragments, the collision energy was in the range 15–40 eV. Argon was used as the collision gas.

## Results

An impure novel AKR protein preparation recovered from affinity chromatography [1] was further separated by 2-D gel electrophoresis yielding sets of multiple spots at five positions (fig. 1). Excision of material from these locations, followed by in-gel digestion and MALDI MS and ESI tandem MS, identified the major material (spots marked 1a–i in fig. 1) as the novel AKR protein, while the major contaminant (spots marked 2a, b in fig. 1) was found to be fructose-bisphosphate aldolase, and a minor contaminant (spot 3, fig. 1) was found to be pyruvate kinase. The remaining spot regions corresponded to multimeric forms of the AKR preparations (dimeric, spot 4; trimeric, spot 5, fig. 1). The dense cluster of major protein spots (segment 1 in fig. 1) corresponds to a molecular weight of 36.5 kDa and a pI of 7.6, values which fit well with those predicted from previous AKR analysis [1]. MALDI MS of peptide extracts from all visible spots in segment 1 of figure 1 showed that they contained the same protein species, and this was also confirmed by tandem MS. A large portion (62%) of the AKR sequence could be analyzed in this approach (fig. 2). The amino acid sequence determined was in accordance with that deduced from cDNA, and established an AsnAsn segment at positions 8 and 9 of the AKR sequence, previously not fully defined. The MS analysis also showed that the methionine at position 1 in the sequence deduced from the cDNA was not present. Instead, we found that the N-terminal amino acid is alanine modified by acetylation after removal of the initiator methionine residue in the nascent protein (fig. 2). In addition, 15 other segments were determined (legend to fig. 2), verifying pre-

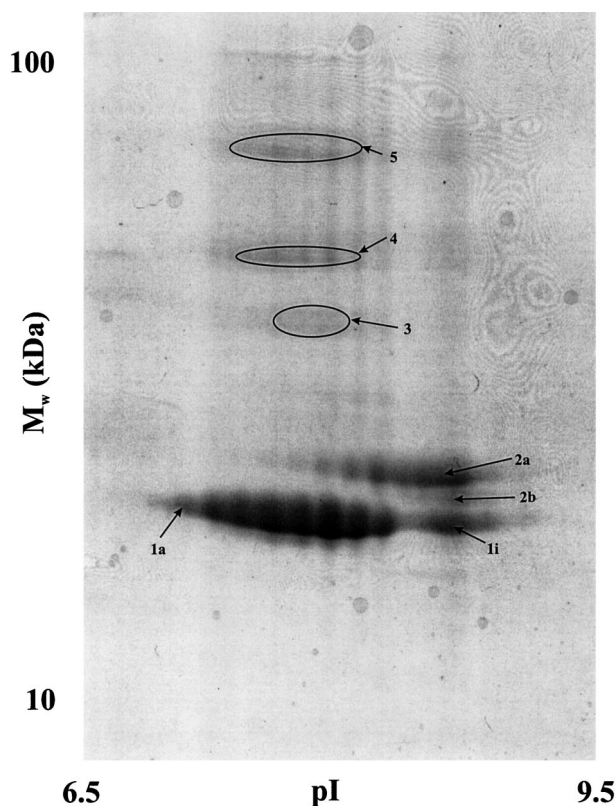


Figure 1. Two-dimensional gel electrophoresis of the AKR protein sample recovered from affinity chromatography and carboxymethylated [1]. Protein (5  $\mu$ g) was loaded and the gel was stained with Coomassie brilliant blue. Horizontally, isoelectric focusing; vertically, SDS/PAGE; pI values and approximate molecular weights are indicated. Spots: 1 a–i, AKR; 2 a, b, fructose-bisphosphate aldolase; 3, pyruvate kinase; 4, dimeric form of AKR; 5, trimeric form of AKR.

vious analyses [1]. These results complete the AKR protein primary structure analysis and establish this novel AKR protein to have a 316-residue protein chain with an acetylated alanine N terminus. The other spots detected in the 2-D gel separation were found to contain other proteins, coeluting with the novel AKR from the current affinity step [1], necessitating use of the gel electrophoretic separation step now employed. Interestingly, the AKR protein seems to be prone to aggregation into a dimeric and also a trimeric form, since spots in the high-molecular-weight region (fig. 1, spots 4 and 5) revealed sequence data matching the AKR structure.

## Discussion

Using 2-D gel separation and subsequent MALDI MS and ESI tandem MS, a novel but impure ADH preparation was separated into its constituent proteins. This ‘proteomics’ approach allowed all components to be analyzed, and the novel ADH could be defined as belonging to a third type of protein family with ethanol dehydrogenase (and retinal re-

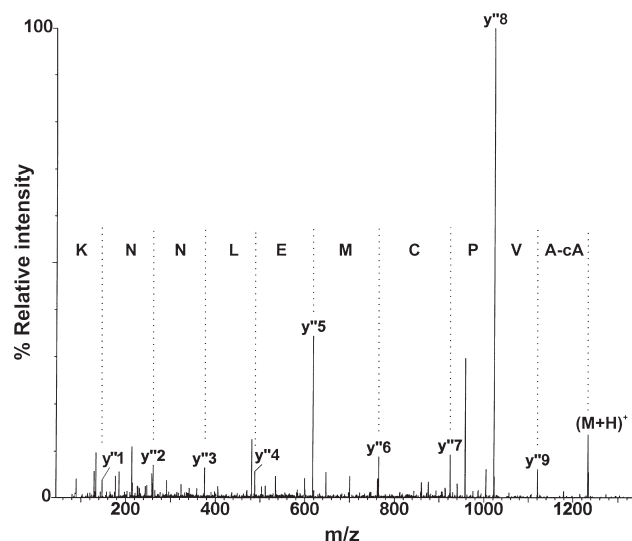


Figure 2. Collision-induced dissociation mass spectrum for the acetylated N-terminal tryptic peptide, revealing the structure to be Acetyl-AVPCMELNNK. Annotations according to the Roepstorff and Fohlman [9] nomenclature, with all y fragments identified as indicated by the sequence interpretation shown. In addition to this N-terminal peptide, the following tryptic peptides were analyzed by tandem mass spectrometry, in all cases supporting the cDNA- and protein sequence-determined structures: M13–K27, H34–R41, H42–K60, R70–K78, L79–K86, S96–K117, A118–K128, A156–R169, L170–K177, G20–K222, P223–K233, T244–R251, F252–K263, E272–K283, E285–R294.

ductase) activity in vertebrate tissues. The new protein is an AKR enzyme and starts with an acetylated alanine residue in a 316-residue polypeptide chain. A tendency to polymerization is also demonstrated, while the multiplicity of charges in the major set of protein spots (area 1, fig. 1) is interpreted to represent partly oxidized carboxymethylated and deamidated protein species.

The family assignment of a vertebrate alcohol dehydrogenase (retinal reductase) as an AKR protein increases the complexity of the ADH enzyme system considerably. MDR, SDR, and AKR are now the three protein families to which major enzyme components in vertebrate tissues assign. Coupled with genetic data assembled from the recently available human genome, the present results show a remarkable multiplicity in the oxidation of alcohols and retinoids. Many MDR classes and the present AKR form constitute enzymes with ethanol dehydrogenase activity, while additional SDR forms also exhibit ethanol, retinol, or steroid ADH activities [9]. Combined, the multiplicity gives protection against losses of function, and suggests that the basic ADH function in both alcohol detoxication and retinoid interconversion is biologically important.

Methodologically, the present results demonstrate that proteomics using 2-D gel separations and MS is also efficient in conventional enzyme characterization.

*Acknowledgements.* This work was supported by grants from the Swedish Research Council (03X-3532, 03X-10832, K5104-20005891), the Emil and Wera Cornell Foundation, the Swedish Society of Medicine, and the European Commission (EC BIO4 CT97 2123).

- 1 Crosas B., Cederlund E., Torres D., Jörnvall H., Farrés J. and Parés X. (2001) A vertebrate aldo-keto reductase active with retinoids and ethanol. *J. Biol. Chem.* **276**: 19132–19140
- 2 Jörnvall H., Höög J.-O. and Persson B. (1999) SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett.* **445**: 261–264
- 3 Peralba J. M., Cederlund E., Crosas B., Moreno A., Julià P., Martínez E. et al. (1999) Structural and enzymatic properties of a gastric NADP(H)-dependent alcohol dehydrogenase: conclusions on origin and function of the vertebrate class IV enzyme. *J. Biol. Chem.* **274**: 26021–26026
- 4 Jörnvall H., Persson B., Krook M., Atrian S., González-Duarte R., Jeffery J. et al. (1995) Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**: 6003–6013
- 5 Napoli J. L. (1999) Interactions of retinoid binding proteins and enzymes in retinoid metabolism. *Biochim. Biophys. Acta* **1440**: 139–162
- 6 Duester G. (2000) Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *Eur. J. Biochem.* **267**: 4315–4324
- 7 Jez J. M. and Penning T. M. (2001) The aldo-keto reductase (AKR) super family: an update. *Chem. Biol. Interact.* **130–132**: 499–525
- 8 Jörnvall H. and Jollès P. (2000) Protein structure analysis of today: proteomics in functional genomics. In: proteomics in Functional Genomics, pp. XI–XIII, Jollès P. and Jörnvall H. (eds), Birkhäuser, Basel
- 9 Roepstorff P. and Fohlman J. (1984) Proposal for a common nomenclature of sequence ions in mass spectra of peptides. *Bio-med. Mass Spectrom.* **11**: 601



To access this journal online:

<http://www.birkhauser.ch>

---