Volume 220, number 1, 209-213 FEB 05000 August 1987

Aldose reductase and ρ-crystallin belong to the same protein superfamily as aldehyde reductase

Deborah Carper, Chihiro Nishimura, Toshimichi Shinohara, Bernard Dietzchold*, Graeme Wistow, Cheryl Craft+, Peter Kador and Jin H. Kinoshita

*National Eye Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, *Wistar Institute, 36th at Spruce St, Philadelphia, PA 19104* and *+Dept of Psychiatry, University of Texas Health Science Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75235, USA*

Received 30 May 1987

Aldose reductase (EC 1.1.1.21) has been implicated in a variety of diabetic complications. Here we present the first primary sequence data for the rat lens enzyme, obtained by amino acid and cDNA analysis. We have found structural similarities with another NADPH-dependent oxidoreductase: human liver aldehyde reductase (EC 1.1.1.2). The identity between these two enzymes is 50%. Both enzymes share approx. 40–50% homology with ρ-crystallin, a major lens protein present only in the frog, *Rana pipiens*. We propose that aldose reductase, aldehyde reductase and ρ-crystallin are members of a superfamily of related proteins.

Aldose reductase; Aldehyde reductase; ρ-Crystallin; Structural homology; cDNA sequence

1. INTRODUCTION

Aldose reductase (AR; EC 1.1.1.21) catalyzes the reduction of a variety of sugars to sugar alcohols. In diabetes and galactosemia, increased AR activity leads to high levels of sorbitol and galactitol, respectively, in the cells of many tissues [1]. Accumulation of sugar alcohols has been shown to cause osmotic cataracts in the lens [2–4]. Aldose reductase is also thought to play a key role in diabetic complications of three other target tissues, namely, nerve, kidney and retina [5–9]. Currently, a variety of AR inhibitors are being tested as a possible new treatment modality for diabetics, although side effects are always a concern.

In order to study the function and tissue specificity of AR and expand on drug designs for AR inhibitors, we initiated a study of the structure of AR using polypeptide and DNA sequencing. While obtaining the sequence for AR, we found considerable homology with another NADPH-dependent reductase: aldehyde reductase (EC 1.1.1.2). In addition, both reductases were found to share homologies with frog lens ϱ-crystallin, adding to the number of crystallins known to be related to enzymes [10–11].

In this article, we present sequence data which compare AR, aldehyde reductase and ϱ-crystallin and discuss the functional and evolutionary relationships of these proteins. In addition, we present the partial nucleotide and amino acid sequence for AR.

Correspondence address: D. Carper, National Eye Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA

2. MATERIALS AND METHODS

N-terminal partial peptide sequence was determined as described [12] using tryptic and cyanogen bromide-cleaved peptide fragments from purified rat lens AR [13] (see fig.1). Oligonucleotide probes designed from this sequence were used to screen a λgt11 rat lens cDNA library (Clonetech, Palo Alto,

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

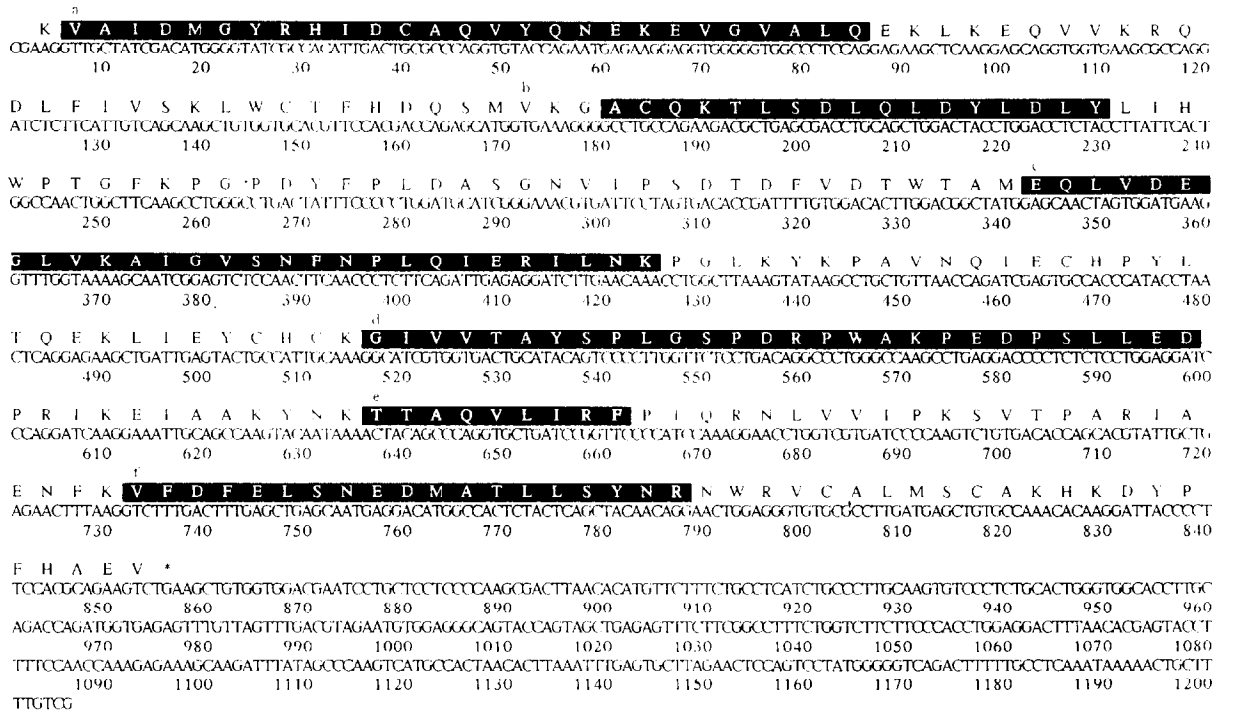

Fig.1. The sequence of the 1206 bp cDNA clone for rat lens AR. The upper line shows the predicted amino acid sequence with the positions of peptides determined by amino acid sequencing indicated in black. Peptides a, d, e and f are tryptic peptide fragments and peptides b and c are CNBr-cleaved fragments. The first three amino acids of peptide b could not be determined by protein sequence. The lower line shows the nucleotide sequence and is numbered. We found two discrepancies between the peptide sequence and the predicted amino acid sequence. The amino acids obtained from the peptide sequence that were different are: S at codon position 40 and W at position 66.

CA). Three oligonucleotides were made using two different approaches [14]. One oligo (20-mer) was made with all possible codons (oligo 696 = TAT-(C)CAA(G)AAT(C)GAA(G)AAA(G)TGGGT), resulting in a mixture of 32 different combinations. The other two oligos were much longer (35- and 36-mer) and had only one combination based on codon usage tables [15]. The sequences are: oligo 697 = TCCGCCCAGGTGTACCAGAACGAGA-AGTGGGTGGG and oligo 698 = CTGGACT-ACCTGGACCTGTACCTGATGCACTGGCCC. The first approach dilutes the probe but guarantees the correct combination. In the latter approach, the uncertainty of each codon is not as critical, since increased probe length confers probe specificity. Indeed, comparison of our final sequence shows that both probes were about 90% correct. Statistically, a 36-mer probe should be 85% correct for specific hybridization [14]. Some of the codon error was due to discrepancies in our initial amino acid sequence. For oligo 697, the codon TCC for serine should have been TGC (cysteine); the codon TGG for tryptophan should have been GAG (glutamic acid). For oligo 698, the codon ATG for methionine should have been ATT (isoleucine). The sequence for oligo 698 was based partially on highly conserved sequences from aldehyde reductase and ϱ-crystallin.

The cDNA clone 10Q was subcloned into the *Eco*RI site of M13 and sequenced using the dideoxy chain termination method of Sanger (as described in the BRL M13 cloning/dideoxy sequencing manual). In addition to the universal M13 primer, synthetic primers designed from the sequence were used. These primers were synthesized using an Applied Biosystems DNA synthesizer (model 360B). The cDNA insert was completely sequenced in both directions.

```
          10          20          30          40          50
AASCVLLHTGQKMPLIGLGTWKSEPGQVKAAVKYALSVGYRHIDCAAIYGNEPEIGEAL-      ALD
                                                KVAIDMGYRHIDCAQVYQNEKEVGVALQ      AR

    60          70          80          90          100         110
---KEDVGPGKAVPREELFVTSKLWNTKHHPEDVEPALRKTLADLQLEYLDLYLMHWPYA      ALD
EKLKEQV-----VKRQDLFIVSKLWCTFHDQSMVKGACQKTLSDLQLDYLDLYLIHWPTG      AR
                                                LERSLRDVGMDYLDLFLMHWPVS      RHO

    120         130         140         150         160         170
FERGDNPFPKNADGT-ICYDSTHYKETWKALEALVAKGLVQALGLSNFNSRQIDDILSV      ALD
FKPGPDYFPLDASGNVIPSD-TDFVDTWTAMEQLVDEGLVKAIGVSNFNPLQIERILNK      AR
LKPSGASDPSDKDKPFI-YDNVDLCATWEALEARKDAGLVRSLGVSNFNRRQLERILNK      RHO

        180         190         200         210         220
ASER--PAVLQVECHPYLAQNELIAHCQARGLEVTAY-PLGSS-DRAWRDPDEPVLLEEP      ALD
PGLKYKPAVNQIECHPYLTQEKLIEYCHCKGIVVTAYSPLGSP-DRPWAKPEDPSLLEDP      AR
PGLKYKPVCNQVECHVYLNQNKLHSYCKSKDIVLVTYSVLGSHRDRNWVDLSLPVLLDDP      RHO

        240         250         260         270         280
VVLALAEKYGRSPAQILLRWQVQRKVICIPKSITPSRILQNIKVFDFTFSPEEMKQLNAL      ALD
RIKEIAAKYNKTTAQVLIRFPIQRNLVVIPKSVTPARIAENFKVFDFELSNEDMATLLSY      AR
ILNKVAAKYNRTSAEIAMRFILQKGIVVLAKSFTPARIKQNLGVFEFELKPEDMKSLESL      RHO

        300         310         320
NKNWRYIVPMLTVDGKRVPRDAGHPLYPFNDPY                                  ALD
NRNWRVCALMSC---------AKHKDYPFHAEV                                  AR
DRNLHYG-PFREV--------KQHPEYPFHDEY                                  RHO
```

Fig.2. Comparison of the amino acid sequences of human liver aldehyde reductase (ALD), rat lens aldose reductase (AR) and frog lens ϱ-crystallin (RHO). Dots represent matches between AR and aldehyde reductase or between AR and ϱ-crystallin. Some variation in the aldehyde reductase sequence has been observed by Wermuth et al. [18]. Position 177 is either E or V and position 209 is either Y or C. The former choices are derived from protein sequencing.

3. RESULTS AND DISCUSSION

Using the oligonucleotide probes designed from the partial amino acid sequence, 15 cDNA clones were isolated from a rat lens λgt11 library. One of these, 10Q, gave hybridization with both probes 697 and 698. Probe 696 is a shorter version of 697 and was not used in this particular screening. The cDNA insert from 10Q was isolated and completely sequenced (fig.1). The insert is 1206 bp in length with an open reading frame covering the first 854 bases and encoding 284 amino acids (or a molecular size of approx. 32300 Da). This corresponds to about 85% of the full size expected for

AR. All of the AR peptides which were sequenced can be accounted for in the 284 amino acids encoded by the open reading frame, indicating that 10Q encodes rat lens AR. A polyadenylation signal is observed at position 1187, suggesting that the 3′-untranslated region of the cDNA is virtually complete.

The partial primary structure of AR was compared with the protein data bank of the National Biomedical Research Foundation. We found a significant similarity (50%) between frog lens ϱ-crystallin and the sequence for AR. The ϱ-crystallin protein sequence deduced from a partial cDNA sequence is about ¾ of the full size [16,17]. We also discovered that rat lens AR has 50% identity with human liver aldehyde reductase, which was recently sequenced in its entirety at the amino acid level [18]. There was slightly less similarity between aldehyde reductase and ϱ-crystallin (43%). The similarities among these three proteins are shown in fig.2; the numbering is according to the complete sequence of aldehyde reductase. Highly conserved areas can be seen. Local identities as high as 84% were observed, in spite of different species and different tissues. For example, there are 16 out of 19 amino acids in common between AR and aldehyde reductase beginning at position 96. Another striking area of similarity exists between AR and ϱ-crystallin beginning at position 169, where 13 amino acids in a row are conserved between these two proteins.

This degree of similarity clearly suggests that all three proteins belong to the same superfamily with related structures and evolutionary origins. The relationship between AR and aldehyde reductase is perhaps not surprising, since both enzymes exhibit overlapping substrate specificity and similar function. Both have been implicated in models of the secondary complications seen in diabetes, with the potential to elevate sugar alcohol levels.

Recently, the active sites for both aldehyde reductase and AR have been reported [19]. The sites are homologous and include a tetrapeptide with the sequence I-P-K-S. Our sequence data show that this tetrapeptide is part of a highly conserved carboxy-terminal region, beginning at position 259. The tetrapeptide sequence is only partially conserved in ϱ-crystallin.

The relationship between ϱ-crystallin, a major structural component of the lens in the genus *Rana pipiens*, and a superfamily of oxidoreductases, is consistent with other observations that so-called 'taxon-specific' crystallins are very closely related to normal cellular enzymes. ϵ-Crystallin in the duck appears to be identical to lactate dehydrogenase-B4 in terms of sequence and enzyme activity [10]. Other close relationships have been observed between other crystallins and enzymes [11]. AR activity was measured in the frog lens, but only low levels were detected. ϱ-Crystallin may lack enzyme activity altogether, or it may have a quite different substrate specificity; in either case this may be reflected in the sequence divergence at the proposed active site.

The roles of AR and aldehyde reductase in normal metabolism and under conditions of hyperglycemia are currently under intense investigation. Several potential therapeutic agents designed as AR inhibitors are under development and a detailed understanding of the structure, function and expression of this enzyme is required. The sequence data presented here should provide useful tools for this endeavor.

REFERENCES

[1] Kador, P.F., Robison, W.G. jr and Kinoshita, J.H. (1985) Annu. Rev. Pharmacol. Toxicol. 25, 691–714.
[2] Gabbay, K.H. and Kinoshita, J.H. (1972) Isr. J. Med. Sci. 8, 1557–1561.
[3] Kinoshita, J.H., Fukushi, S., Kador, P. and Merola, L.D. (1979) Metabolism 28 (suppl. 1), 461–469.
[4] Kinoshita, J.H., Kador, P. and Datiles, M. (1981) J. Am. Med. Assoc. 246, 257–261.
[5] Yue, D.K., Hanwell, M.A., Satchell, P.M. and Turtle, J.R. (1982) Diabetes 31, 789–794.
[6] Judzewitsch, R.G. et al. (1983) N. Engl. J. Med. 308, 119–125.
[7] Beyer-Mears, A., Ku, L. and Cohen, M.P. (1984) Diabetes 33, 604–607.
[8] Robison, W.G. jr, Kador, P.F. and Kinoshita, J.H. (1983) Science 221, 1177–1179.
[9] Frank, R.N., Kern, R.J., Kennedy, A. and Frank, K.W. (1983) Invest. Ophthal. Vis. Sci. 24, 1519–1528.
[10] Wistow, G.J., Mulders, J.W.M. and DeJong, W.W. (1987) Nature 326, 622–624.
[11] Wistow, G. and Piatigorsky, J. (1987) Science 236, 1554–1556.

[12] Shinohara, T. et al. (1987) Proc. Natl. Acad. Sci. USA 84, in press.
[13] Shiono, T., Sato, S., Reddy, V.N., Kador, P.F. and Kinoshita, J.H. (1987) in: Enzymology and Molecular Biology of Carbonyl Metabolism, vol. 232, pp. 317–324.
[14] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[15] Grantham, R., Gautier, C. and Gouy, M. (1980) Nucleic Acids Res. 8, 1893–1912.
[16] Tomarev, S.I., Zinovieva, R.D., Dolgilevich, S.M., Luchin, S.V., Krayer, A.S., Skryabin, K.G. and Gause, G.G. jr (1984) FEBS Lett. 171, 297–302.
[17] Gause, G.G. jr, Tomarev, S.I., Zinovieva, R.D., Aruthyunyan, K.G. and Dolgilevich, S.M. (1986) in: The Lens: Transparency and Cataract, pp. 171–179.
[18] Wermuth et al. (1987) in: Enzymology and Molecular Biology of Carbonyl Metabolism 232, pp. 297–307, Alan R. Liss, New York.
[19] Mojana, N.A., Lyons, C. and Flynn, T.G. (1987) Fed. Proc. 46, 2154.