Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 309 (2007) 18–31
DEVELOPMENTAL BIOLOGY
www.elsevier.com/developmentalbiology

Three male germline-specific aldolase A isozymes are generated by alternative splicing and retrotransposition

Soumya A. Vemuganti [a,b], Timothy A. Bell [a,c], Cameron O. Scarlett [f,1], Carol E. Parker [f], Fernando Pardo-Manuel de Villena [a,c,d,e], Deborah A. O'Brien [a,b,d,*]

[a] *Laboratories for Reproductive Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[b] *Department of Cell and Developmental Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[c] *Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[d] *Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[e] *Carolina Center for Genome Sciences, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[f] *UNC-Duke Michael Hooker Proteomics Center, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*

Received for publication 27 April 2007; revised 9 June 2007; accepted 12 June 2007
Available online 18 June 2007

Abstract

Enzymes in the glycolytic pathway of mammalian sperm are modified extensively and are localized in the flagellum, where several are tightly bound to the fibrous sheath. This study provides the first evidence for three novel aldolase isozymes in mouse sperm, two encoded by *Aldoart1* and *Aldoart2* retrogenes on different chromosomes and another by *Aldoa_v2*, a splice variant of *Aldoa*. Phylogenetic analyses and comparative genomics indicate that the retrogenes and splice variant have remained functional and have been under positive selection for millions of years. Their expression is restricted to the male germline and is tightly regulated at both transcriptional and translational levels. All three isozymes are present only in spermatids and sperm and have distinctive features that may be important for localization in the flagellum and/or altered metabolic regulation. Both ALDOART1 and ALDOA_V2 have unusual N-terminal extensions not found in other aldolases. The N-terminal extension of ALDOA_V2 is highly conserved in mammals, suggesting a conserved function in sperm. We hypothesize that the N-terminal extensions are responsible for localizing components of the glycolytic pathway to the fibrous sheath and that this localization is required to provide sufficient ATP along the length of the flagellum to support sperm motility.

Keywords: Spermatogenesis; Sperm; Glycolysis; Aldolase isozymes; Retrogene; Retrotransposition; Alternative splicing; Fibrous sheath; LC-MALDI-MS

Introduction

Mammalian spermatogenesis includes three phases: a mitotic proliferation period that expands the number of spermatogonia, a prolonged meiotic prophase allowing spermatocytes to undergo recombination followed by two meiotic divisions, and a post-meiotic period where haploid spermatids differentiate into highly polarized sperm that are specialized to achieve fertilization. The program of gene expression that directs this developmental process has several distinct features and produces a large number of transcripts that are restricted to spermatogenic cells (DeJong, 2006; Eddy, 2002; Elliott and Grellscheid, 2006). Microarray analyses estimate that at least 4% of the mouse genome produces male germ cell-specific transcripts, predominantly during the meiotic and post-meiotic phases of spermatogenesis (Pang et al., 2006; Schultz et al., 2003; Shima et al., 2004). Essential processes that occur during these two phases include the generation and subsequent packaging of the haploid genome within an extremely condensed nucleus, formation of specialized sperm structures such as the acrosome and flagellum, organization of surface domains essential for fertilization, and the development of complex signaling and metabolic cascades

* Corresponding author. Department of Cell and Developmental Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. Fax: +1 919 966 1856.
E-mail address: dao@med.unc.edu (D.A. O'Brien).
[1] Present address: Analytical Instrumentation Center, School of Pharmacy, University of Wisconsin, 777 Highland Avenue, Madison, Wisconsin 53705, USA.

0012-1606/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2007.06.010

that regulate sperm function and gamete interaction. Multiple mechanisms contribute to this extensive diversification of gene function including gene duplication, retrotransposition and alternative splicing. Many conserved pathways, as well as germ cell-specific structures and processes, have unique variants with restricted expression during spermatogenesis.

Glycolysis is an important conserved pathway that has been modified substantially during mammalian spermatogenesis. This central metabolic pathway is composed of ten enzymes that convert glucose to pyruvate, with a net production of 2 ATPs per glucose molecule. Pyruvate can be further metabolized in the mitochondria via the Krebs cycle and oxidative phosphorylation, or alternatively converted to lactate by lactate dehydrogenase (LDH). Eight glycolytic enzymes and LDH have multiple isoforms with distinct patterns of expression in different tissues (Steinke et al., 2006). At least three of these gene families include members with restricted expression in germ cells. Two germ cell-specific isozymes with distinct enzymatic properties, glyceraldehyde 3-phosphate dehydrogenase-S (GAPDHS) and lactate dehydrogenase C (LDHC), are encoded by intron-containing genes that arose by gene duplication. In the mouse *Gapdhs* is expressed only in spermatids (Bunch et al., 1998; Welch et al., 1992), while *Ldhc* is expressed in spermatogenic cells and oocytes (Coonrod et al., 2006; Goldberg, 1977; Li et al., 1989). Phosphoglycerate kinase-2 (*Pgk2*), which encodes another germ cell-specific isozyme, is an intronless gene that evolved by retrotransposition (i.e., a retrogene) from the *Pgk1* gene (Boer et al., 1987; McCarrey and Thomas, 1987). Although *Pgk2* is transcribed in primary spermatocytes, the PGK2 protein is translated only in post-meiotic spermatids (Bluthmann et al., 1982; Vandeberg et al., 1981).

Other glycolytic enzymes, including hexokinase 1 (Mori et al., 1993, 1998), phosphoglucose isomerase (Buehr and McLaren, 1981), phosphofructokinase 1 (Yamada et al., 2004), aldolase (Gillis and Tamblyn, 1984) and enolase (Edwards and Grootegoed, 1983; Gitlits et al., 2000), have unique structural or functional properties in spermatogenic cells and sperm. The molecular basis for these distinctive properties has not been determined in most cases. One exception is HK1-S, the hexokinase 1 isozyme that is derived from alternative splicing and lacks the porin-binding domain responsible for binding to mitochondria (Mori et al., 1993, 1998).

Energy production in sperm is compartmentalized in distinct regions of the flagellum, with mitochondria and oxidative phosphorylation restricted to the middle piece and glycolysis localized in the longest segment known as the principal piece (Eddy et al., 2003). Although there are differences between species, mammalian sperm typically exhibit a high rate of glycolysis that is correlated with motility (Hoskins, 1973; Mann and Lutwak-Mann, 1981). Multiple studies indicate that glycolysis provides a significant proportion of ATP in both mouse and human sperm (Mukai and Okuno, 2004; Peterson and Freund, 1969; Williams and Ford, 2001). Gene targeting studies provide compelling evidence that glycolysis in spermatozoa (Miki et al., 2004) rather than mitochondrial ATP production (Narisawa et al., 2002) is essential for maintaining sperm motility and male fertility in the mouse. Distinctive features of sperm glycolytic enzymes may be important for localization in the principal piece and/or altered regulation of this key metabolic pathway.

Several sperm glycolytic enzymes are difficult to solubilize because they are tightly bound to the fibrous sheath, a cytoskeletal structure that defines the limits of the principal piece in the sperm flagellum. We found that GAPDHS, aldolase A (ALDOA), lactate dehydrogenase A (LDHA), and pyruvate kinase remain attached to the fibrous sheath throughout a rigorous isolation procedure (Bunch et al., 1998; Krisfalusi et al., 2006). GAPDHS is larger than other GAPDH family members and contains a novel proline-rich extension at the N-terminus (Welch et al., 1992, 2000) that may mediate binding to the fibrous sheath. Our proteomic and immunoblot analyses identified two ALDOA bands in mouse sperm, with the larger 50,000 molecular weight band always present in isolated fibrous sheaths (Krisfalusi et al., 2006). ALDOA, along with several other glycolytic enzymes, was also identified in fibrous sheaths isolated from human sperm (Kim et al., 2006). Consistent with this localization, an earlier study found that 90% of aldolase activity in bovine sperm could not be solubilized with detergents (Gillis and Tamblyn, 1984). That study also determined that sperm and muscle aldolases had distinct kinetic properties.

The fructose-1,6-bisphosphate aldolase gene family in vertebrates contains three well-characterized members: *Aldoa*, which is ubiquitously expressed with particularly high levels in the muscle, aldolase B (*Aldob*) which is expressed at high levels in the liver and kidney, and aldolase C (*Aldoc*) which is predominantly expressed in the nervous system (Penhoet et al., 1966, 1969; Rutter et al., 1968). All three of these isozymes have molecular weights of ~39,000, significantly smaller than the isoform detected in fibrous sheath. To determine the identity and origin of this larger ALDOA isoform, we used genomic, molecular and proteomic methods to examine all the aldolase variants expressed during mouse spermatogenesis and present in mature sperm. Our analyses identified three spermatogenic cell-specific aldolase isoforms in mouse, two encoded by retrogenes and a third resulting from alternative splicing of the *Aldoa* gene.

Materials and methods

Genomic analyses to identify Aldoa-related sequences

Ensembl Mouse BlastView (http://www.ensembl.org/Mus_musculus/blastview) was used to search the mouse genome (NCBI database, mouse build 34) for sequences highly related to the mRNA sequence of mouse *Aldoa* (accession number NM_007438). Chromosomal regions containing significant matches were aligned to the mRNA sequence of *Aldoa* using ClustalW (http://www.ebi.ac.uk/clustalw/). All chromosome numbers included refer to Assembly: NCBIm36, December 2005; Genebuild: Ensembl, April 2006; Database version: 42.36c. For each sequence, we identified all insertions and deletions resulting in a shift in open reading frame (ORF) and internal stop codons. We assume that sequences containing such insertions/deletions or stop codons are pseudogenes that do not encode functional aldolase enzymes. Intronless sequences that contain a full-length ORF with conserved start and stop codons were classified as putative retrogenes. These included sequences found on chromosome 4 and chromosome 12, now termed *Aldoart1* and *Aldoart2*. The Mouse Genome Informatics database (http://www.informatics.jax.org/) previously identified both of these regions as *Aldoa* pseudogenes. *Aldoart1* contains an extended ORF capable of encoding a protein with an additional 55 amino acids at the N-terminus. Further genomic comparisons identified a homologous