Developmental Biology **250,** 24–43 (2002)
doi:10.1006/dbio.2002.0769

The *t* Complex Distorter 2 Candidate Gene, *Dnahc8*, Encodes at Least Two Testis-Specific Axonemal Dynein Heavy Chains That Differ Extensively at Their Amino and Carboxyl Termini

Sadhana A. Samant,* Olugbemiga Ogunkua,* Ling Hui,* John Fossella,† and Stephen H. Pilder*,[1]

**Department of Anatomy and Cell Biology, Temple University School of Medicine, Philadelphia, Pennsylvania 19140; and †The Sackler Institute for Developmental Psychobiology, Cornell University, New York, New York 10021*

Homozygosity for the *t* haplotype allele of the testis-specifically expressed axonemal dynein heavy chain (axDHC) gene, *Dnahc8*, has been linked to male sterility resulting from aberrant sperm motility. However, the near absence of *Dnahc8* expression has been associated with male sterility resulting from an early breakdown in sperm flagellar development. Although axDHCs are integral participants in flagellar motility, a role in flagellar morphogenesis has never been attributed to a member of this highly conserved gene family. To gain a better understanding of this presumed novel role for *Dnahc8*, we have studied the organization and expression of full-length *Dnahc8*$^{+}$ and *Dnahc8*^{t} transcripts. Our results demonstrate the existence of at least two alternatively spliced, testis-specific *Dnahc8* mRNAs transcribed from both the + and *t* alleles. A highly expressed isoform encodes a protein with significant homology nearly throughout to the γ heavy chain of the *Chlamydomonas* axonemal outer arm dynein, while a more poorly expressed isoform codes for a protein whose sequence diverges significantly from that of other axDHCs at both its N and C termini. While *in situ* hybridization studies demonstrate that both mRNA species accumulate exclusively in mid to late spermatocytes, each isoform shows spatial independence. Additional experiments demonstrate the existence of a testis-expressed mRNA with no significant open reading frame, a portion of which is antisense to the 5′-untranslated region of the highly divergent *Dnahc8* isoform. The cumulative data imply that *Dnahc8* may have acquired functional plasticity in the testis through the tightly controlled expression of both typical and unusual isoforms.

***Key Words:* t complex; *t* haplotype; dynein heavy chain; stem domain; motor domain; sperm tail; flagellar motility; flagellar morphogenesis; axoneme; fibrous sheath.**

INTRODUCTION

A *t* haplotype (*t*) is a member of a family of naturally occurring, structure–function variants of the *t* complex, the proximal third of *Mus musculus* chromosome (Chr) 17. One feature that distinguishes a *t*-containing homolog from its wild type (+) counterpart is the presence of four, large, nonoverlapping inversion polymorphisms spanning nearly the entire *t* complex region (Hammer *et al.*, 1989; Fig. 1A), resulting in almost complete recombination suppression in heterozygous (+/*t*) animals over a distance of ~20 cM. In addition, +/*t* males transmit their *t* homolog to nearly all their progeny, a trait known as transmission ratio distortion (TRD), while males homozygous for *t* are sterile without exception.

The biochemical mechanisms underlying +/*t*-TRD and *t*/*t*-associated male sterility are as yet undetermined. Nevertheless, previous studies have suggested that the basis of both phenotypes is similar if not identical (Lyon, 1984, 1986), with TRD resulting from an interaction between three or more distorter/sterility "factors" (hereafter referred to as distorters) and a recently isolated and partially characterized responder gene product (*Tcr*; Herrmann *et al.*,

[1] To whom correspondence should be addressed. Fax: (215) 707-2966. E-mail: stephen.pilder@temple.edu.

0012-1606/02 $35.00

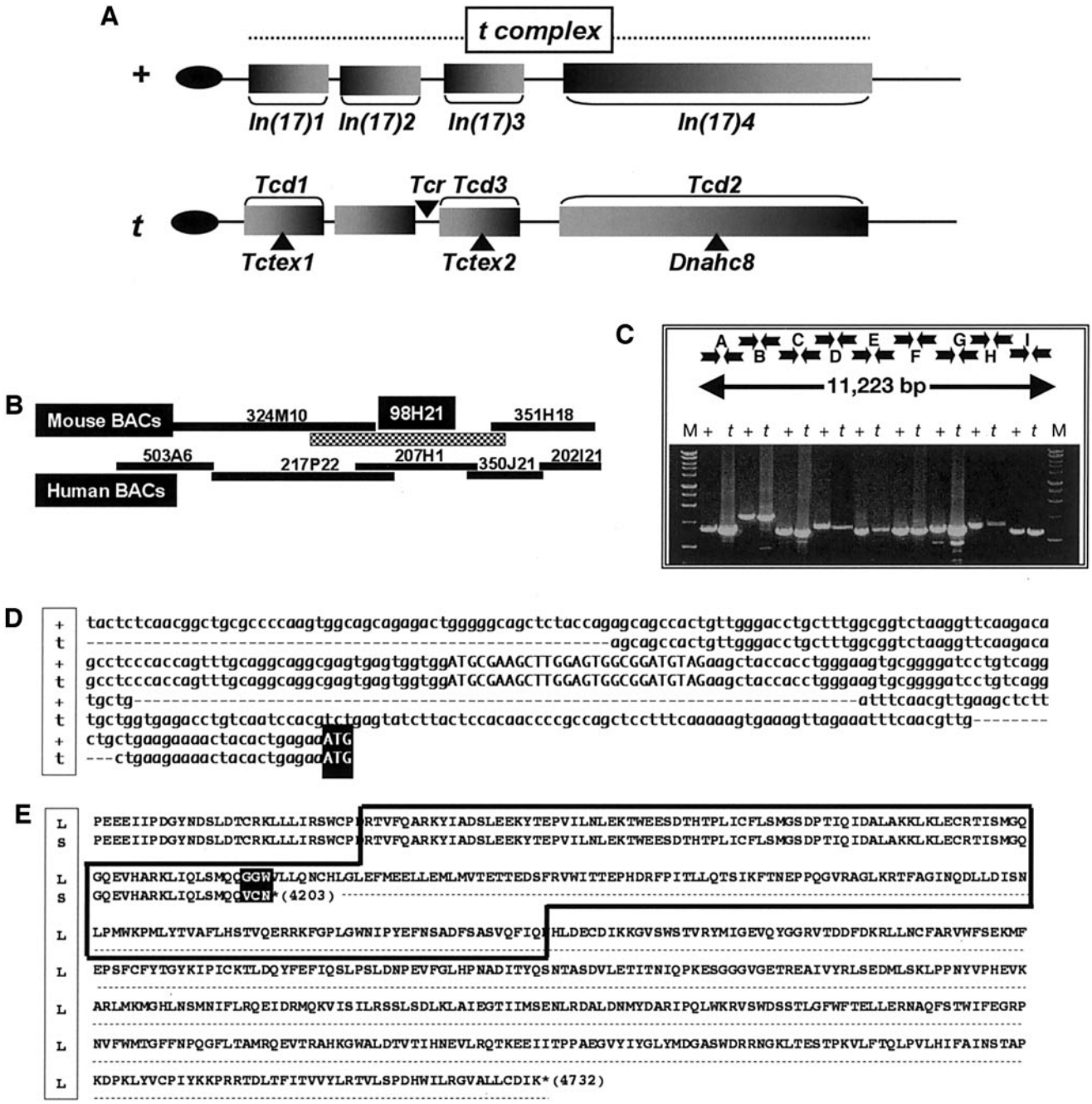

FIG. 1. (A) The + and *t* homologs of the mouse *t* complex. The wild type (+) and *t* haplotype (*t*) homologs of the mouse *t* complex are represented by horizontal lines with four differentially shaded rectangles superimposed upon each line. The filled ellipse on the left end of each homolog represents the centromere of that homolog; the rectangles, labeled *In(17)1–In(17)4* below the + homolog, correspond to large (~1-cM to ~10-cM) +–*t* relative inversions, and the gross positions of the distorters (*Tcd1*, *Tcd3*, and *Tcd2* from proximal to distal) are shown above the *t* homolog, as is the position of the previously isolated responder gene (*Tcr*). The positions of dynein light chains *Tctex1* and *Tctex2* and the axonemal dynein heavy chain *Dnahc8* are shown below the first, third, and fourth inversions, respectively, on the *t* homolog. (B) Mouse and human BACs containing *Dnahc8* orthologous sequences are aligned. In mouse, only the checkerboard BAC, 98H21, has been completely sequenced, while 324M10 has been partially sequenced. All human BACs have been completely sequenced (see text for accession numbers). (C) Nine pairs of oppositely oriented arrows are shown (top) representing primer pairs derived by aligning the sequence of 98H21 with the sequences of three of the five human BACs (shown in B), and identifying highly conserved regions (putative *Dnahc8* exons). These primer pairs, designated A–I, were used to amplify overlapping + and *t* cDNA fragments (see agarose gel at bottom) covering the central 75% (11,223 bp, see line below primer pairs) of *Dnahc8* mRNAs by RT-PCR of total + and *t* testis RNAs. (D) Alignment

1999), and sterility caused by homozygosity of the distorters (particularly *t complex distorter 2* or *Tcd2*) acting in a responder allele-independent manner (Fig. 1A).

Several avenues of investigation have been explored in attempts to elucidate the molecular basis of *t*-associated male phenotypes. Numerous sperm function analyses have demonstrated that the motility of sperm from +/*t* and *t*/*t* animals is both quantitatively and qualitatively abnormal (Olds-Clarke and Johnson, 1993; Pilder *et al.*, 1993; Redkar *et al.*, 2000). Molecular genetic studies of the *t* complex have established that the dynein light chains (DLCs) *Tctex1* and *Tctex2* localize to the inversions containing *t* complex distorters 1 and 3 (*Tcd1* and *Tcd3*), respectively, and both genes have been shown to encode alternative + and *t* proteins that are present in mouse sperm (Harrison *et al.*, 1998). Additionally, an axonemal dynein heavy chain (axDHC) gene, *Dnahc8*, is centrally located within the large *Tcd2* locus (Fig. 1A) and may be testis-specifically expressed (Fossella *et al.*, 2000). Interestingly, males heterozygous for *t* and any of numerous overlapping Chr 17 polymorphisms containing a nearly null allele of *Dnahc8* (also referred to as *Dnahc8*^{s}) are sterile and produce sperm that display flagellar curvature phenotypes very similar to "curlicue," the signature trait exhibited by sperm from *t*/*t* homozygotes (Olds-Clarke and Johnson, 1993; Pilder *et al.*, 1993; Redkar *et al.*, 1998; Fossella *et al.*, 2000). These diverse findings have been incorporated into a simple, yet compelling hypothesis which proposes that the *t*-distorter alleles code for defective components of one or more axonemal dyneins, the microtubule- (MT-) associated, minus end-directed (retrograde) molecular motors responsible for flagellar propulsion in eukaryotes (Harrison *et al.*, 1998; Herrmann *et al.*, 1999; King, 2000b; Fossella *et al.*, 2000).

Not surprisingly, males homozygous (or compoundly heterozygous) for *Dnahc8*^{s}-containing chromosomes are also sterile due to sperm immotility. However, lack of sperm movement in this case does not result from the simple loss of axonemal dynein arms, as is observed in typical cases of primary ciliary dyskinesia (PCD), but from a severely disrupted, sperm flagellar developmental program ("whipless" phenotype; Pilder *et al.*, 1993; Phillips *et al.*, 1993; Samant *et al.*, 1999; Fossella *et al.*, 2000). "Whipless" flagella contain grossly distorted axonemes (or no axonemes at all), unassembled fibrous sheaths, and ballooned periaxonemal compartments, reminiscent of the most severe, sperm-specific forms of Dysplasia of the Fibrous Sheath (DFS), a heritable human male sterility condition of unknown genetic origin(s) (Chemes *et al.*, 1987; Rawe *et al.*, 2001).

The expression of the DFS-like "whipless" phenotype in testes that are nearly null for *Dnahc8* expression raises the intriguing possibility that *Dnahc8* plays an important role in mammalian sperm tail morphogenesis. One line of evidence supporting the plausibility of this idea is the recent demonstration that a cytoplasmic dynein heavy chain (cytDHC) gene, *DHC1b*, plays a retrograde transport role in the flagellar assembly process of *Chlamydomonas reinhardtii* (Pazour *et al.*, 1999; Porter *et al.*, 1999). Nevertheless, all previous studies of axonemal dyneins imply that, unlike cytoplasmic dyneins, the sole function of these complex, highly conserved molecular motors is to generate sliding force between adjacent axonemal MT doublets, thus producing flagellar/ciliary movement (Witman, 1992; Dutcher, 1995; Gibbons, 1996; Porter, 1996; Supp *et al.*, 1997; Pazour *et al.*, 1999; Porter *et al.*, 1999; Olbrich *et al.*, 2002; Ibanez-Tallon *et al.*, 2002).

In the present report, we attempt to clarify the function of *Dnahc8* by analyzing its molecular organization and expression pattern. We demonstrate that, in the adult mouse, *Dnahc8* is a testis-specific, meiotically expressed gene that may have gained functional flexibility through its ability to generate multiple mRNA isoforms, at least one of which contains terminal modifications entirely novel for an axDHC.

MATERIALS AND METHODS

Sequences and Computer-Assisted Sequence Analysis

Sequence analysis tools and design programs used in this study were obtained from the following Web sites: www.bionavigator.com; www.ncbi.nlm.nih.gov; www.expasy.ch; http://restools.sdsc.edu; and www.ensembl.org/Mus_musculus. All sequences have been submitted to GenBank and either have been assigned an accession number or an accession number is pending.

Mouse Dnahc8 Gene Cloning and Sequencing

In a previous study, we identified a mouse BAC clone, 98H21 (AF342999), that harbored the ESTs *Hst6.7b* and *Ploop1* from the axDHC gene *Dnahc8* (Fossella *et al.*, 2000). This BAC, with an insert of ~150 kb, was sequenced in its entirety at the Nucleic Acid/Protein Research Core (NAPCORE) facility (The Children's Hospital of Philadelphia, Philadelphia, PA). Nucleic acid search and alignment programs, such as BLAST and CLUSTALW (accurate), were employed to extract orthologous sequences to 98H21 from public databases. The alignment results and various exon-finding programs (Genscan, Genewise, Grail) were used to derive

of 5′ RACE products (5′ UTRs only) of + and *t* cDNA alleles of *Dnahc8*. The first in-frame ATG of each sequence is shown at its 3′ end, indicated by white capitalized letters with a black background. A short, upstream, out-of-frame open reading frame (uORF) in each allele is indicated by black capitalized letters. Sequences not present in one allele or the other are represented by dashed lines. (E) Alignment of the long (L) and short (S) C termini of DNAHC8 as deduced from translating 3′ RACE products. The sixth AAA domain (RTV. . . IQN) is boxed.

putative exons within the 98H21 genomic sequence. By assembling these exons, an ~11-kb putative partial mouse *Dnahc8* cDNA sequence was constructed. This sequence was used to select primer pairs for amplifying 1- to 2-kb-long overlapping cDNA fragments. (A table of all primer sequences used in the present study is available upon request from S.H.P.) RT-PCR was subsequently performed on mouse testis mRNAs from *129Sv*-+/+ (+/+) and F1[*C57BL/6* × *129Sv*]-t^{w32}/t^{w5} (*t*/*t*). To minimize the PCR error rate, Herculase Enhanced Polymerase Blend (Stratagene, La Jolla, CA) was used in all RT-PCR assays, and assays were performed multiple times from different mRNA preparations. RT-PCR products were cloned in a ZeroBlunt Topo PCR vector (Invitrogen, Carlsbad, CA) and sequenced. To obtain cDNA 5′- and 3′-end sequences, RACE reactions were first performed by using the commercially available Sure-RACE mouse multitissue cDNA panel from NIH *Swiss Webster* mice (Origene Technologies Inc., Rockville, MD) or Marathon-ready testis cDNA from 9- to 11-week-old *BALB/c* mice (Clontech, Palo Alto, CA). Sequence data obtained from these preliminary experiments were used to select gene-specific primers and to perform PCR or RT-PCR of +/+ or *t*/*t* mouse testis cDNA or mRNA templates. Additional 5′ and 3′ RACE reactions were performed with the SMART RACE cDNA amplification kit (Clontech) with total RNA extracted from +/+ or *t*/*t* mouse testes. 5′- and 3′-cDNA sequences were also extracted by screening mouse testis cDNA libraries, constructed with poly(A)$^+$ mRNA from +/+ or *t*/*t* adult mouse testes, using both oligo(dT)- and random hexamer-priming with SuperScript reverse transcriptase (GIBCO/BRL, Rockville, MD). *Eco*RI-adapted cDNAs were ligated to λZAP Express vector arms and packaged with Gigapack III packaging extracts (Stratagene). ^{32}P-labeled probes used for library screening included the A cDNA fragment (see Fig. 1C, and Results) for the 5′ end, and both *u3L* and *u3S* (see Fig. 4, and Results), unique sequences from long or short 3′-ended *Dnahc8* isoforms, respectively. Library screening was as described (Pilder *et al.*, 1992). Labeling of probes, hybridizations, and posthybridization washes were essentially done as described elsewhere (Pilder *et al.*, 1991). Genomic sequence information around the 5′ untranslated region (UTR) of the highly divergent *Dnahc8* mRNA isoform was obtained by sequencing part of the BAC clone 324M10, which overlaps 98H21 proximally (see above, and Samant *et al.*, 1999).

Northern/RNA Dot Blot Analyses

Total RNAs (2–10 μg) from adult mouse testis tissue (+/+, *t*/*t*, and *Dnahc8*^{s}/*Dnahc8*^{s} in *Mus musculus* genetic background) were electrophoresed through 0.7–0.8% agarose–3%formaldehyde denaturing gels in 1× RNA Borate buffer (3.24 mM sodium tetraborate, 30.8 mM boric acid, and 0.2 mM EDTA, pH 8.0). RNA samples were added to an equal volume of 2× RNA–borate loading buffer (65.6% formamide, 7.9% formaldehyde, 2.4× RNA–borate buffer, 80 μg/ml ethidium bromide, 1 mg/ml each of bromophenol blue and xylene cyanole), incubated at 65°C for 10 min, and quickly chilled on ice before loading. For the "long" Northern gel (Fig. 4C), RNAs were separated for ~20 h at 4°C with one change of running buffer. Gels were then processed for blotting and probing as previously described (Fossella *et al.*, 2000). For tissue-specificity experiments, commercially available mouse multiple tissue Northern blots and RNA dot blots (Origene) were used.

PCRs of cDNA Panels

PCR analysis of cDNAs from 24 mouse tissues was carried out to detect trace amounts of *Dnahc8* transcripts with the mouse RapidScan Gene Expression panel (Origene). Primer pairs for two different *Dnahc8*-ESTs were used: one from the first *P-Loop* region of *Dnahc8* and the other from the *Hst6.7b* stem region (Fossella *et al.*, 2000). For the 564-bp-long *P-Loop*-specific product, PCR was carried out with (Forward) 5′-CTGATAGATGCTACATCACGC-3′ and (Reverse) 5′-GTAAGCTGCTCTTCGCAGAG-3′ primers. For the *Hst6.7b*-specific 394-bp-long product, PCR was performed by using (Forward) 5′-ATCACCGACTCCGCTAATG-3′ and (Reverse) 5′-AAACTTCAAAAGACTTCTCCCC-3′ primers. The cDNA panel was normalized by PCR with human *β-actin* primers as recommended by the manufacturer.

In Situ Hybridizations

Localization of *Dnahc8* transcripts was carried out by *in situ* hybridizations primarily on paraformaldehyde-fixed, paraffin-embedded testis tissue thick sections from adult mouse (Novagen, WI). In addition, other ciliated tissues, such as lung epithelium and eye, as well as embryo sections of 8-, 10-, 12-, 14-, and 16-day-old mice (NIH *Swiss Webster* strain) were tested. All antisense and corresponding sense riboprobes were synthesized with the DIG-labeling and detection kit (Roche Molecular Biochemicals/Enzo Diagnostics) and the RNA polymerases (T3, T7, or SP6) from Ambion (TX). *In situ* hybridization and detection procedures were essentially as described by Wilkinson (1995). Results were visualized by brightfield/differential interference contrast (DIC) microscopy (Nikon Eclipse E800 microscope) and photographed by using the Optronics (DEI-750) imaging system.

RESULTS

Isolation of Dnahc8$^+$ and Dnahc8^t cDNA clones. We sequenced an ~150-kb mouse bacterial artificial chromosome clone (BAC clone 98H21; GenBank AF342999) previously shown to hybridize to two small *Dnahc8* ESTs (*Hst6.7b* and *P1-Loop*; Fossella *et al.*, 2000; Fig. 1B). BLAST searches of the nonredundant (nr) nucleotide database with 98H21 query sequences revealed significant homology to the central ~75% of various DHCs (data not shown). Foremost among these were the γ heavy chain of the *C. reinhardtii* axonemal outer dynein arm (GenBank CRU15303) and the genomic sequence of part of a putative human *Dnahc8* ortholog, consisting of the central three of five overlapping BAC clones (GenBank AL391415, AL035555, AL034345, AL353700, and AL035690 from 5′ to 3′; Fig. 1B). The human ortholog of *Dnahc8* (*DNAH8*) mapped to the short arm of human Chr 6 between bands p21.1 and p21.31 in conserved synteny with genes near *Dnahc8* on mouse Chr 17 (Trachtulec and Forejt, 2001).

Alignment programs (see Materials and Methods) were used to identify highly conserved regions of 98H21 and the orthologous human BACs. Together with results generated by computer-based exon identification analysis of 98H21, the highly conserved regions provided a potential exon contig covering an estimated 75% (the central portion) of the human and mouse *Dnahc8* mRNAs. The sequence of this contig was subsequently employed to develop sets of oligonucleotide primer pairs to amplify overlapping fragments of mouse *Dnahc8* cDNAs via RT-PCR of testis

mRNAs (Fig. 1C). cDNAs were completed at their 5′ and 3′ ends by RACE (rapid amplification of cDNA ends) from uncloned, synthetically anchored testis cDNAs.

Initial 5′-RACE experiments generated one *Dnahc8*$^{+}$ 5′ end and a similar but nonidentical *Dnahc8*^{t} 5′ end. The *Dnahc8*$^{+}$ 5′ end contained 55 bases upstream of the *Dnahc8*^{t} 5′ end, but included an internal deletion of 76 bases and a more distal 11-base insertion relative to the *Dnahc8*^{t} sequence upstream of the putative translational start site (Fig. 1D). Three prime-RACE produced two alternative 3′ ends from wild type testis cDNA, one ~1200 bases longer than the other, and three alternative 3′ ends from *t/t* cDNA, a long one like its wild type counterpart, and two short ends, both similar to the single short 3′ end in wild type (one slightly shorter than the other).

Sequence analysis demonstrated that short and long 3′-RACE ends (3S and 3L, respectively) were identical for their 5′-121 bp, much of this overlapping the cDNA fragment from which the forward nested 3′-RACE primers had been derived. As a result of sequence divergence immediately distal to the 121st base, dynein heavy chains translated from the *Dnahc8* mRNAs with 3S ends (GenBank AF356521, AF356523, AF363577) terminated 3 amino acids distal to the site of divergence, 529 amino acid residues N-terminal of a protein translated from *Dnahc8* mRNAs with the 3L ends (GenBank AF356520, AF356522; Fig. 1E). Thus, DHCs encoded by the messages with the 3S ends contained motor units that were missing the distal half of their 6th AAA (ATPases associated with cellular activities) domains and all of their C-terminal non-AAA domains (King, 2000a; Mocz and Gibbons, 2001).

The mRNAs encoding proteins with abbreviated C termini contained 3′ UTRs of 596 bases or 446 bases, approximately three- to fourfold the length of the 138-base 3′ UTR of the messages encoding the protein with the nontruncated C terminus (Fig. 2A). The 3′ UTR of the mRNAs with 3S ends contained two canonical poly(A) addition signals, AAUAAA, commencing 165 and 79 bases from the termini of the single wild type 3S end and the longer of the two *t* 3S ends, respectively, with the most 5′-AAUAAA apparently utilized as a poly(A) addition signal by the shortest (446 base) *t* 3S end. A more degenerate sequence, AAUAAGAAA, starting 20 bases upstream of the single wild type 3S end and the longer of the two *t* 3S ends, provided a poly(A) addition signal near the end of the 596-base 3′ UTRs. The + and *t* 3L ends contained only a slightly noncanonical poly(A) addition signal, AUUAAA, starting 24 bases upstream of poly(A) addition.

Comparison of the cDNA and 98H21 sequences revealed that the BAC contained 68 central exons from the 3L-ended mRNAs, while additional sequence at the 5′ and 3′ ends of these mRNAs mapped by Southern blot analysis to more proximal and distal overlapping BACs (324M10 and 351H18), respectively (Fig. 2B). However, the cDNAs with 3S ends contained the first 66 exons from 98H21 as well as additional sequence from the 5′-overlapping BAC, 324M10. At the end of the 66th 98H21 exon (bp 141,801 of 98H21), these mRNAs proceeded to read through the splice donor site into the intron between 98H21 exon 66 and exon 67, where they terminated (Fig. 2B).

Analysis of Dnahc8 expression. Previous experiments with RNA isolated from a limited number of male mouse tissues suggested that *Dnahc8* expression might be testis-specific (Fossella *et al.*, 2000). To further test this hypothesis, we performed Northern, dot blot, and RT-PCR analyses with probes and primer pairs generated from sequences common to both short and long *Dnahc8* mRNAs (*Dnahc8* ESTs *Hst6.7b* and *P1-Loop*; Fossella *et al.*, 2000) and RNAs isolated from a wide range of tissues, both male and female (Fig. 3A). We also examined expression in several adult ciliated tissues and at numerous embryonic developmental stages by *in situ* hybridization (Fig. 3B). The results of these assays reinforced the earlier conclusion that *Dnahc8* is testis-specifically expressed in the mouse.

In situ hybridization of testis sections with antisense *Hst6.7b* and *P1-Loop* RNA probes revealed that mRNA expression commenced at about stage VI or VII of the seminiferous epithelial cycle in midpachytene primary spermatocytes (Fig. 3C). Signal first appeared as dense, usually (but not always) single granules accompanied by fainter, more peripheral, diffuse staining that proceeded to become more intense in later stage spermatocytes. The diffuse staining was evident until approximately the end of meiosis, but the granular signal seemed to disappear about one or two stages prior to that. No signal was evident in spermatogonia, spermatocytes before approximately midpachytene, spermatids, or nongerm cells, nor was signal apparent in sections probed with the negative control sense strand (Fig. 3C). These data demonstrated that translation of all *Dnahc8* mRNA isoforms must occur prior to spermiogenesis.

Northern blots of testis RNAs from males that were either *Dnahc8*$^{+}$, *Dnahc8*^{t}, or *Dnahc8*^{s} homozygotes were also probed with labeled DNA fragments whose sequences were unique either to 3S (*u3S* probe) or 3L (*u3L* probe) ended mRNAs (Fig. 4A). Both probes hybridized to approximately equal-size mRNA species in *Dnahc8*$^{+}$/*Dnahc8*$^{+}$ and *Dnahc8*^{t}/*Dnahc8*^{t} samples, though the mRNAs with the 3L end appeared to be considerably more abundant than the mRNAs with the 3S end. This difference in abundance could potentially result from the fact that the messages with 3S ends had longer, fairly AU-rich (~60%) 3′ UTRs containing seven/nine AUUU mRNA destabilizing elements, respectively (Shaw and Kamen, 1986), while the shorter 3′ UTRs of the messages with 3L ends shared a lower A+U content (~52%) and no AUUU elements (see Fig. 2A). As expected, no mRNA species in the *Dnahc8*^{s}/*Dnahc8*^{s} mRNA sample was recognized by either probe.

In situ hybridization of adult mouse testis sections with *u3S* and *u3L* antisense RNA probes showed that mRNAs with 3L ends accumulated from about late midpachytene primary spermatocytes to diplotene primary/secondary spermatocytes, and signal was generally intense and predominantly diffuse (Fig. 4B). Interestingly, hybridization

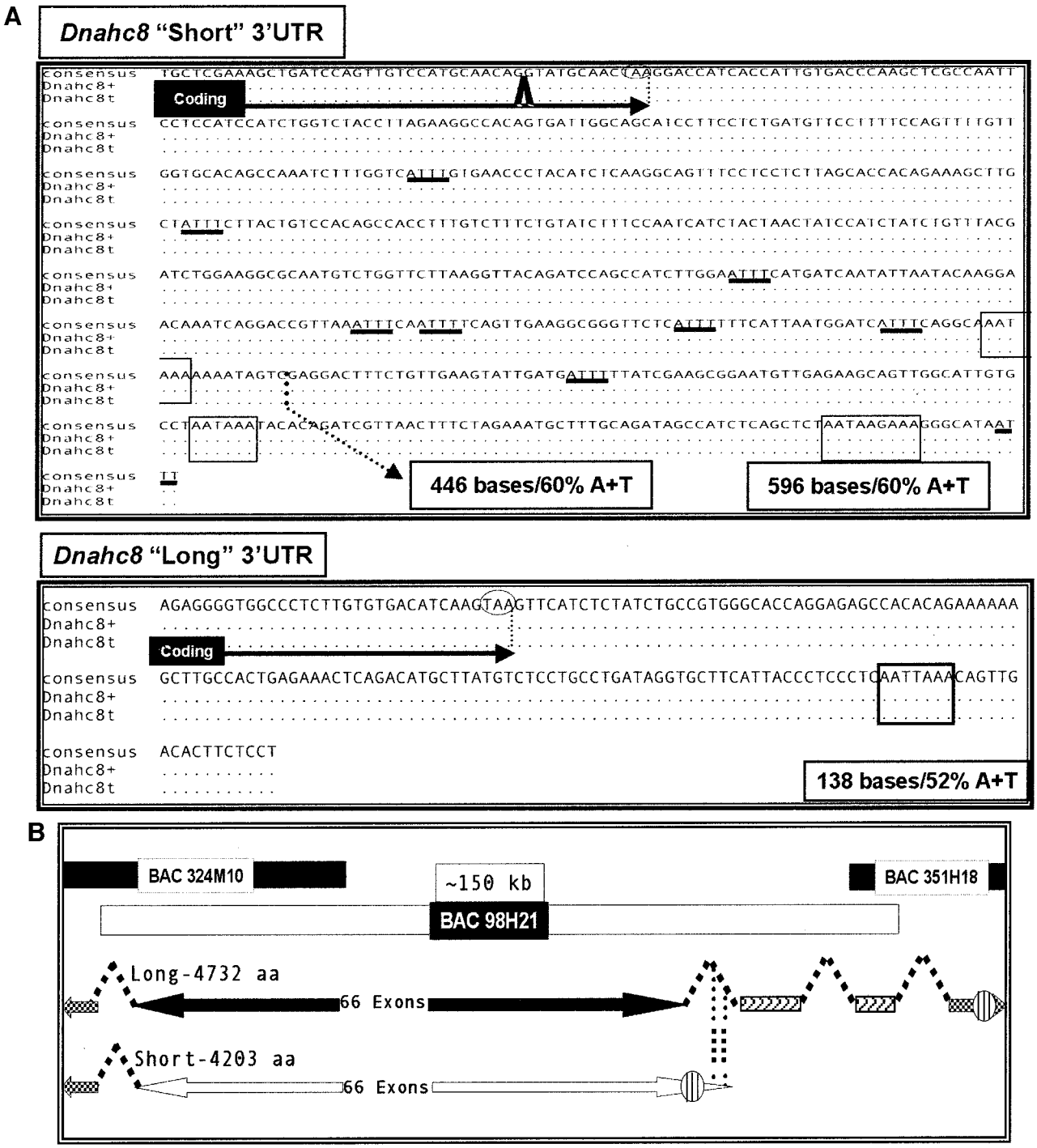

FIG. 2. (A) Two 3′ UTRs of *Dnahc8* as deduced by 3′ RACE. At the top, the 3′ UTR of the "short" *Dnahc8* isoform is shown. At its beginning, the horizontal arrow shows the C-terminal coding region ending at the TAA stop codon (circled). The vertical arrowhead preceding the end of the coding region indicates the site of the GT splice donor not recognized during formation of the short isoform. The underlined ATTT sequences correspond to putative 3′ end mRNA instability signals. Putative poly(A) addition signals are boxed. Below, the 3′ UTR of the "long" *Dnahc8* isoform is shown. At the beginning, the horizontal arrow shows the C-terminal coding region ending at the TAA stop codon (circled). The putative poly(A) addition signal is boxed. There are no ATTT sequences. (B) Formation of the short *Dnahc8* isoform. Sixty-eight of more than 90 exons that form the long *Dnahc8* mRNA isoform are contained within the BAC clone, 98H21. The first 66 exons (and 65 introns) are represented by a solid black horizontal line with arrowheads at each end shown below the white

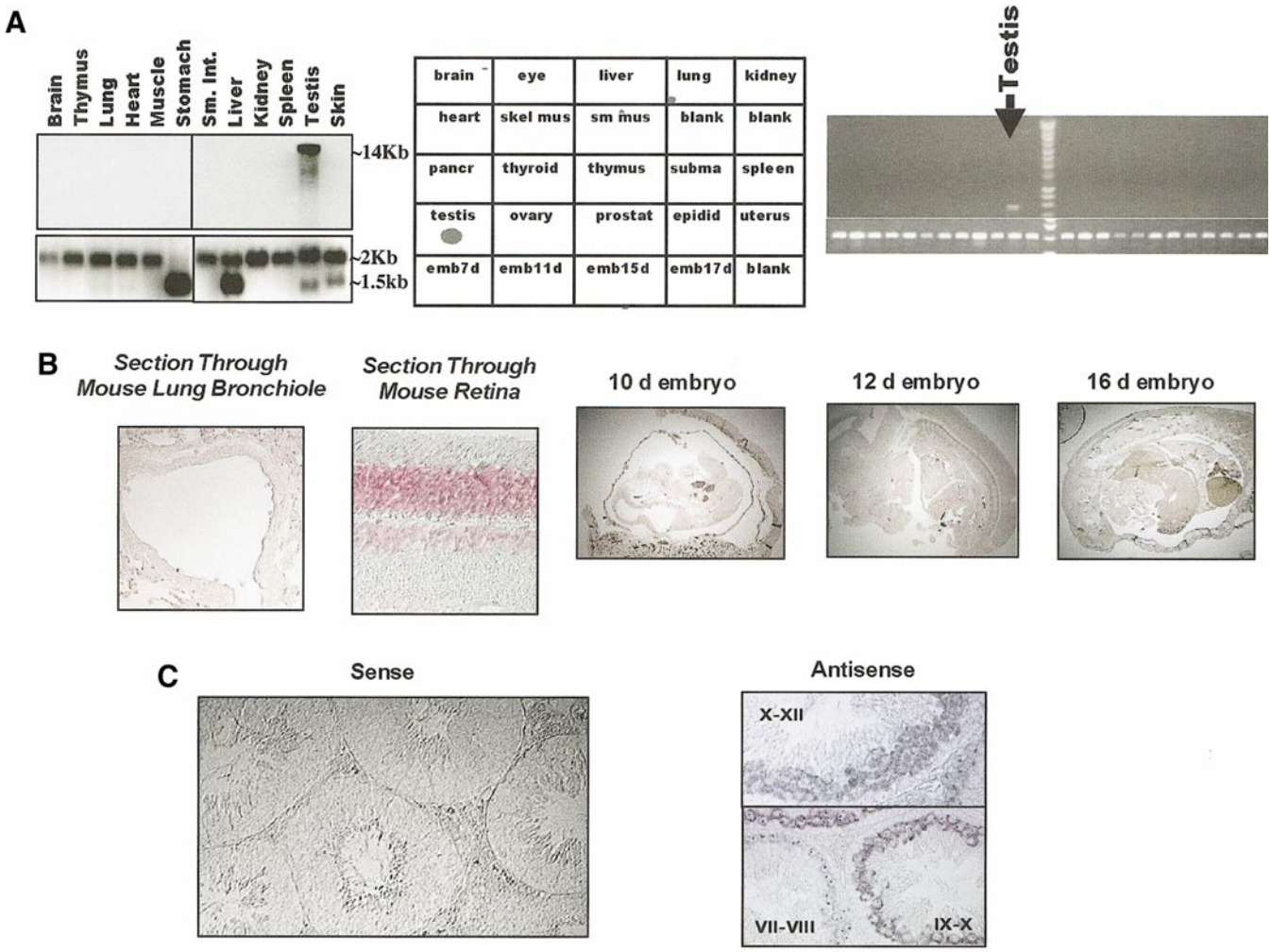

FIG. 3. (A) Results of Northern, RNA dot-blot, and RT-PCR analyses of *Dnahc8* expression in multiple tissues from male and female adults and various embryonic stages. Probes and primer pairs used (in upper panels of Northern blot and RT-PCR analysis) were *Hst6.7b*- and *P1-Loop*-specific (Fossella *et al.*, 2000), two regions from the shared central 75–80% of the stem and motor unit, respectively, of *Dnahc8* mRNAs. Although only the results from probing with the *P1-Loop* are shown, both probes gave identical results. Probe and primer pair used in lower panels of Northern blot and RT-PCR analysis were β-actin-specific. (B) Examples of paraformaldehyde-fixed, paraffin-embedded mouse adult ciliated tissue and embryonic stage-specific thick sections hybridized with *Dnahc8* antisense probes described in (A). Only counterstaining with nuclear fast red is obvious. (C) Results of *in situ* hybridization experiments in which paraformaldehyde-fixed, paraffin-embedded adult mouse testis cross-sections were hybridized with sense (left) and antisense (right) *P1-Loop* probes. Bottom panel of the antisense hybridization shows two different stages stained: primary midpachytene spermatocytes on left (~stages VII-VIII), and slightly later to the right (~stages IX-X). Both stages show differential granular and diffuse staining and differential intensities of each. The top panel shows exclusively diffuse peripheral staining in late primary pachytene, primary diplotene, or secondary spermatocytes (~stages X-XII). Identical results were obtained by hybridizing with *Hst6.7b* antisense probes.

rectangle corresponding to 98H21. The 67th and 68th exons are represented by herringbone patterned horizontal lines separated from the first 66 exons, from each other, and from exons in the distal BAC, 351H18 (checkerboard horizontal line with a right-sided arrowhead), by introns (dashed vertical arrowheads). Exons proximal to those in 98H21 are represented by the checkerboard horizontal line with a left-sided arrowhead and are separated from the 68 exons of 98H21 by an intron (dashed vertical arrowhead) to the left of the solid black horizontal line. The 66 exons (and 65 introns) of the short *Dnahc8* isoform within 98H21 are represented by a white horizontal line with arrowheads at each end shown below the solid black horizontal line and the first right-sided intron. Dashed vertical lines within this intron correspond to poly(A) addition sites of the "short" isoforms. Stop codons in each ORF are represented by circles containing vertical black lines.

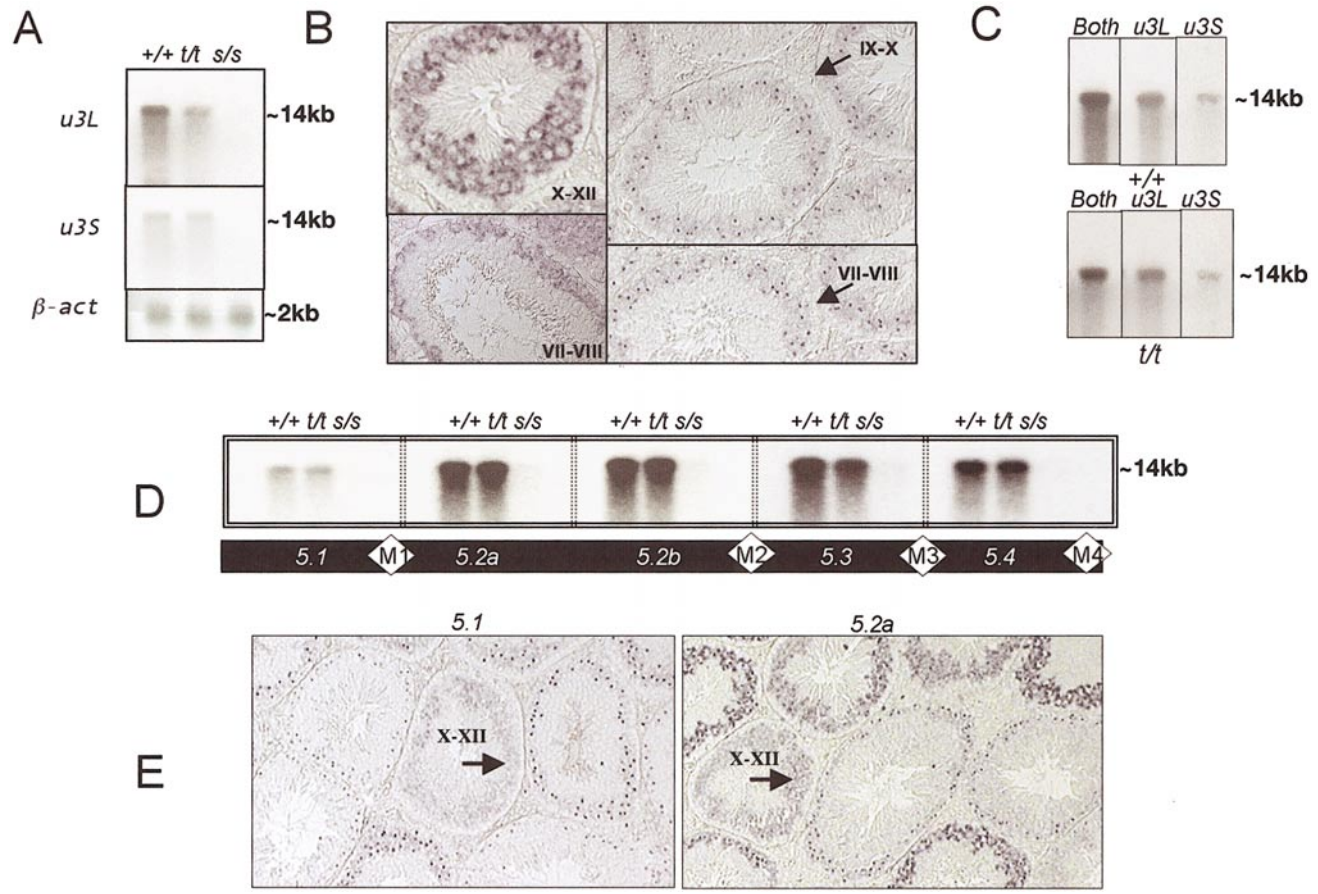

FIG. 4. (A) Northern analysis of *Dnahc8* mRNA levels in the testis of *Dnahc8*$^{+}$/*Dnahc8*$^{+}$, *Dnahc8*^{t}/*Dnahc8*^{t}, or *Dnahc8*^{s}/*Dnahc8*^{s} homozygotes probed with labeled DNA fragments (*u3L*, top; *u3S*, middle) whose sequences were unique either to long or short 3′-ended mRNAs, respectively. Positive control probe: *β-actin* cDNA fragment (bottom). (B) Representative results of *in situ* hybridization experiments in which paraformaldehyde-fixed, paraffin-embedded adult mouse testis cross-sections were hybridized with either *u3L* (left) or *u3S* (right) antisense probes. No signal was evident with sense probes (data not shown). All sections hybridized with the *u3L* probe showed primarily nongranular, intense and diffuse, peripheral staining, while all sections stained with the *u3S* probe showed primarily granular staining with light diffuse peripheral staining. Approximate stages are shown in the figure. Arrows indicate the tubules in question. (C) Northern analysis of *Dnahc8* mRNA levels in the testis of *Dnahc8*$^{+}$/*Dnahc8*$^{+}$ (top) or *Dnahc8*^{t}/*Dnahc8*^{t} (bottom) homozygotes probed with labeled *u3S* (right), *u3L* (left), or both together. mRNAs were separated on a low percentage agarose gel for twice normal distance. (D) Northern analysis of *Dnahc8* mRNA levels in the testis of *Dnahc8*$^{+}$/*Dnahc8*$^{+}$, *Dnahc8*^{t}/*Dnahc8*^{t}, or *Dnahc8*^{s}/*Dnahc8*^{s} homozygotes hybridized with five nearly equal length probes spanning the 5′-most ~1 kb of the *Dnahc8*$^{+}$ 5′ RACE end [*5.1*, *5.2a*, *5.2b*, *5.3*, and *5.4*, from proximal (left) to distal (right)]. The symbols M1–M4 represent the approximate positions of the first four in-frame methionines relative to the five probes. (E) Representative results of *in situ* hybridization experiments in which paraformaldehyde-fixed, paraffin-embedded adult mouse testis cross-sections were hybridized with either labeled antisense *5.1* (left) or *5.2a* (right) probes. Arrows point to stages in which staining by each probe is exclusively nongranular. In both of these tubules, staining appears to be either in late pachytene primary, diplotene primary, or secondary spermatocytes (stages X–XII). No signal was apparent with sense probes (data not shown).

with the *u3S* probe generated the granular signal with very light diffuse, peripheral staining (Fig. 4B). The granular signal appeared to commence and terminate about one to two stages of the seminiferous epithelial cycle prior to the signal produced with the *u3L* probe (from midpachytene primary spermatocytes to late pachytene primary spermatocytes). The intense, predominantly nongranular signal seen previously in primary diplotene/secondary spermatocytes was not evident. However, because the *u3S* probe had been generated by necessity from a 3′ UTR, *in situ* hybridization may not have been completely efficient (Angerer *et al.*, 1987; Angerer and Angerer, 1991).

Additional Northern analyses of testis RNA samples electrophoresed through a low percentage gel for twice

normal distance prior to blotting and probed with *u3S* and/or *u3L* revealed that the mRNAs containing different length 3′ ends appeared to be of equal or nearly equal lengths (Fig. 4C). This result implied that more than one 5′ end existed (a longer one possessed by the mRNAs with 3S ends and a shorter one possessed by the mRNAs with 3L ends), although the equal size of the different mRNA isoforms could have reflected an inability to resolve size differences between very large (≥13.5 kb) mRNAs in our gel system.

Subsequent tests of the possibility that 3S- and 3L-ended mRNAs contained long and short 5′ ends, respectively, were performed. Northern blots containing equal amounts of testis mRNAs from males that were either *Dnahc8*$^{+}$, *Dnahc8*^{t}, or *Dnahc8*^{s} homozygotes were hybridized with five nearly equal length probes spanning the 5′-most ~1.0 kb of the *Dnahc8*$^{+}$ 5′-RACE end (Fig. 4D). The 5′-most probe (*5.1*; generated from sequence upstream of the first in-frame AUG) hybridized to far lower levels of steady state *Dnahc8* mRNA than did any of the four more distal probes (*5.2a*, *5.2b*, *5.3*, and *5.4*; in that order, from 5′ to 3′, with *5.2a* and *5.2b* downstream of the first in-frame AUG but upstream of the second in-frame AUG and overlapping, *5.3* downstream of the second in-frame AUG at its 5′-end and downstream of the third in-frame AUG at its 3′-end, and *5.4* downstream of the third in-frame AUG but just upstream of the fourth in-frame AUG and overlapping *5.3*), despite the fact that *5.1* was the second most GC-rich of the five probes. Thus, the signal strength generated by *5.1* vs those of the other four more distal 5′ probes paralleled the signal strengths produced by *u3S* vs *u3L*, respectively.

In situ hybridizations of adult mouse testis sections with antisense probes nearly identical to either *5.1* or *5.2a* were also performed. The *5.1* probe generated signal comparable to that seen previously with *u3S* (granules, light peripheral staining), except that a few tubules containing late spermatocytes not only displayed a signal, but the signal was predominantly nongranular (diffuse), although weaker in intensity than the signal generated previously when either the *u3L*, *Hst6.7b*, or the *P1-Loop* probe was employed. Signal produced by *5.2a* antisense was identical to that seen in the original *in situ* hybridization experiments (Fig. 3C) when the *Hst6.7b* or *P1-Loop* probe was used (Fig. 4E). When placed in context with our other expression data, these results suggested that a few of the two mRNA isoforms with alternative 3′ ends might possess similar long 5′ ends, while the overwhelming majority of transcripts with the long 3′ end would possess a shorter 5′ end than the great majority of transcripts with the short 3′ end.

To further explore the issue of *Dnahc8* 5′ end heterogeneity, we isolated additional 5′ end clones by probing +/+ and *t/t* testis cDNA libraries prepared by a combination of oligo(dT)- and random-priming with either a cDNA fragment spanning the region from ~750 to ~1650 bp distal to the original 5′-RACE ends or a combination of the *5.1–5.4* fragments. Approximately 90 5′ end-containing plasmids were recently isolated, but only 3 (2 from the *t/t* library and 1 from the +/+ library) have so far been completely sequenced. Analysis of sequenced 5′ ends (Fig. 5) shows that no *t/t* end yet examined initiates as far 5′ as the +/+ 5′ RACE end (+*R1*). In addition, the two *t/t* mRNAs that contain the most 5′ in-frame translational start site (one isolated by RACE [*tR1*] and the other by library screening [*tL1*]) have 5′ UTRs that differ internally from the putative +/+ 5′ UTR as well as from each other. Interestingly, the 5′ UTR of the *tL1* 5′ end is probably too short to hybridize well to the *5.1* probe in either Northern blot or *in situ* hybridization experiments. If similar 5′ ends are abundant in wild type and *t*, this might partially explain the differential levels of *Dnahc8* mRNA hybridization obtained with *5.1* and the other 5′ end probes in both Northern blot and *in situ* hybridization experiments. The third *t/t* 5′ end (*tL2*) initiates 3′ of the 5′-most in-frame translational start site (Fig. 5). If *tL2* is not an artifact of incomplete reverse transcription, at least a portion of the different *Dnahc8* mRNA isoforms could initiate translation from downstream AUG start codons. However, it is not known whether this end is sufficiently short to explain the apparent similarity in size seen on Northern blots between mRNAs that differ by ~1200 bp at their 3′ ends.

As stated above, an additional 5′ end was isolated from the wild type library. Initial analysis suggested that this 5′ end contained novel sequences at both the proximal and distal ends of the 5′ UTR of +*R1*. The distal end of the clone, however, was fused to a testis-expressed Y chromosome repetitive mRNA, but the initial 55 bases of this distal region were not part of the Y-specific message. Sequence analysis of the genomic region surrounding the 5′ UTR of +*R1* revealed that the +*R1* 5′ UTR was extended proximally for 155 bases, distally read through the sequence present in *tR1*, but deleted from +*R1* for an additional 53 bases, then spliced to a new second exon 229 bases downstream where the sequence continued for 55 bases before fusion to the Y-specific element occurred (Figs. 5 and 6A). Additional sequencing of this clone and comparison to the genomic sequence surrounding the *Dnahc8* 5′UTR revealed a third exon of 244 bases beginning 1749 bp upstream of the already extended +*R1* 5′ UTR. Interestingly, all of the splice donors and acceptors between the three new exons (the +*R1* and *tR1* 5′ UTRs were contained within the middle exon) were completely noncanonical unless the construct was viewed in reverse as a complementary strand or antisense mRNA (Fig. 6B). In that context, the sequence revealed an antisense transcript consisting of 3 exons, 721 bases in length, with completely canonical splice donors and acceptors and a canonical AAUAAA polyadenylation signal beginning 20 bases upstream of its 3′ end. This transcript also had no ORF longer than 22 amino acids (Fig. 6B). A subsequent BLAST search of the mouse EST database confirmed our suspicion that this novel *Dnahc8* cDNA 5′ end was most probably part of a testis-expressed *Dnahc8* 5′ UTR-antisense mRNA (see GenBank AA110101). Because this mRNA was not observed with our *5.1* sense strand probe in *in situ* hybridizations or with *5.1* probe hybridiza-

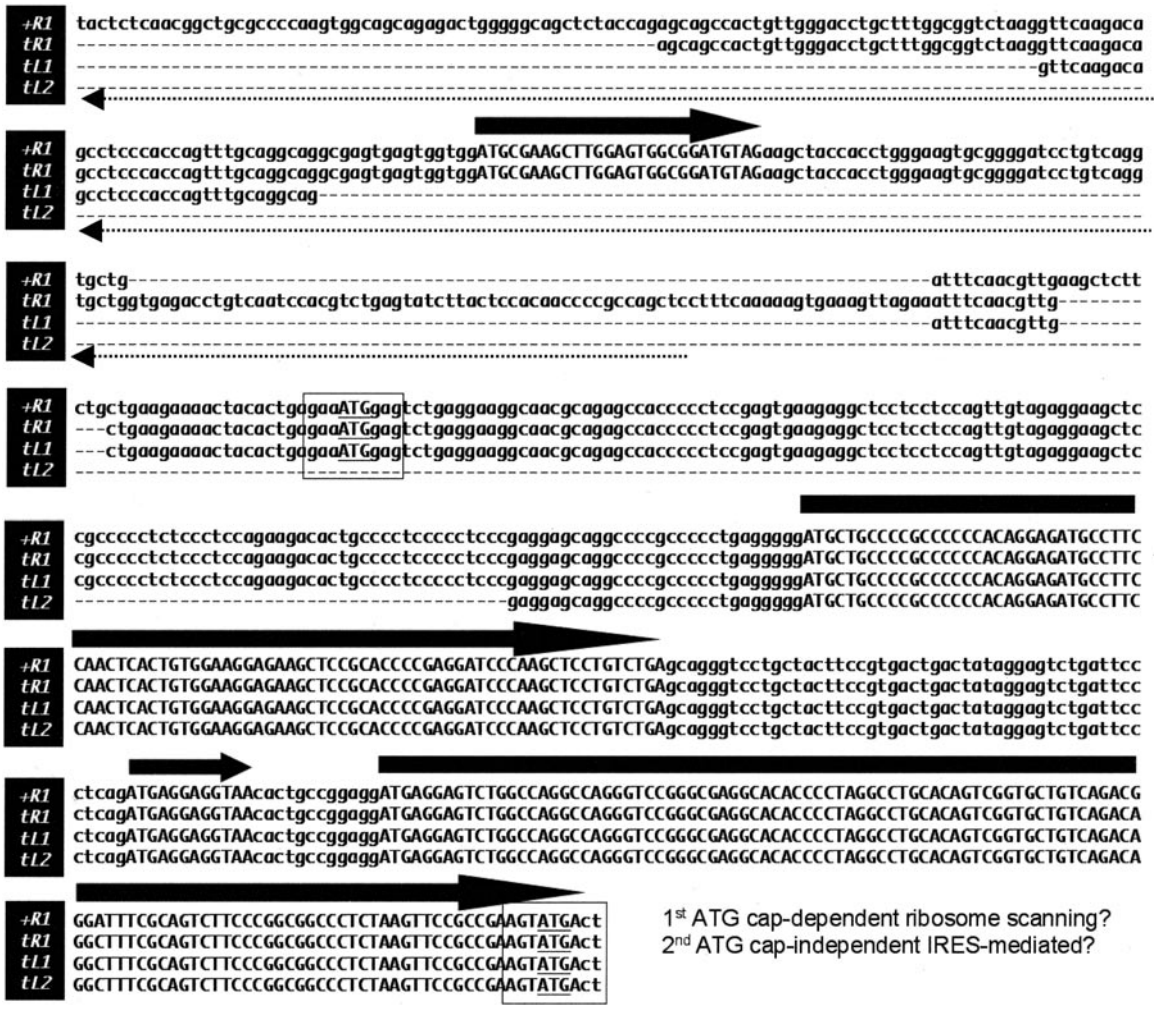

FIG. 5. Alignment of the *Dnahc8*$^{+}$ 5′ end (+*R1*) with three *Dnahc8*^{t} 5′ ends (*tR1*, *tL1*, and *tL2*) up to and including the second in-frame methionine. Both in-frame methionines are underlined and boxed along with flanking sequences that make up the Kozak consensus sequences from −3 to +6. Horizontal arrows above capitalized sequences represent out-of-frame ORFs upstream of each in-frame methionine. Sequences not present in one allele or the other are represented by dashed lines. Dashed lines with left-pointing arrowheads below the 5′ UTR represent part of a testis-expressed complementary antisense mRNA.

tion to Northern blots, we surmised that it was a rarely expressed mRNA. Additional Northern blot analysis with 5× testis RNA loaded on a 2% gel revealed a weak band of approximately 7–900 bp, and a smaller band at about 4–500 bp (data not shown).

Sequence analysis of potential Dnahc8 translational start sites. The most upstream in-frame AUG (translation initiation codon 1 or MET1) in both *Dnahc8*$^{+}$ and *Dnahc8*^{t} is located within a relatively good context for cap-dependent ribosome scanning (CACUGAGAA**AUG**GAG, −3 is a purine, +4 is a G, and + 5 is an A; Gray and Wickens, 1998), although the 5′ UTRs preceding it in +*R1* and *tR1* are relatively long (252 and 262 nucleotides, respectively), each containing an upstream out-of-frame short ORF (uORF), whose initiation codon is in a poor context for cap-dependent ribosome scanning. *tL1*, however, has a 5′ UTR that is internally spliced, producing a relatively short 5′ UTR missing the aforementioned uORF (Fig. 5). Only *tR1* contains stop codons in all three forward reading frames. The other two, +*R1* and *tL1*, contain stop codons in two reading frames, one of those being in frame with MET1. The fourth sequenced 5′ end, *tL2*, initiates downstream of MET1, more than 300 bases upstream of the second in-frame AUG (MET2). The 5′ UTR of *tL2* is burdened by multiple uORFs and out-of-frame AUGs, GC-richness, and an out-of-frame stop codon (UGA), whose first and second bases are the second and third bases of MET2 (Fig. 5). Consequently, if *tL2* is not simply an artifact of incomplete

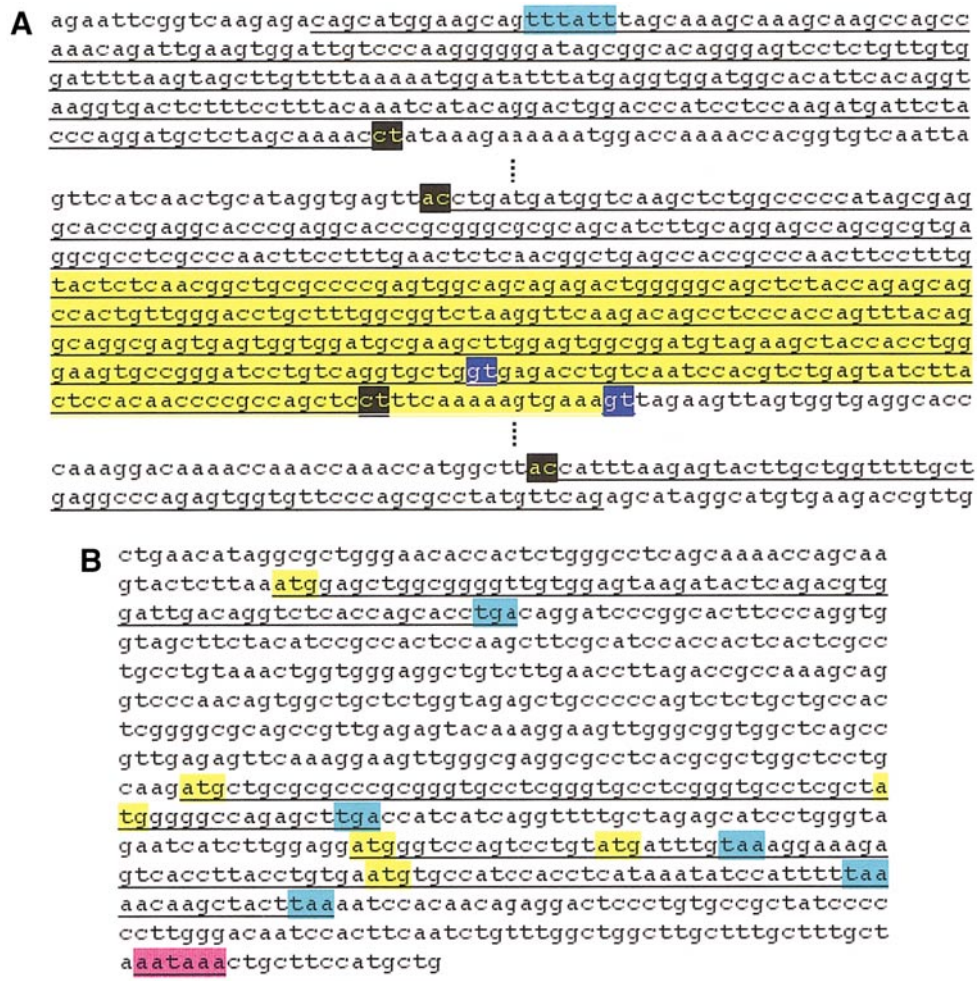

FIG. 6. (A) Genomic sequence surrounding exon 1 (yellow region containing most of 5′ UTRs from *+R1* and *tR1*) of the divergent *Dnahc8* mRNA isoform. Blue boxes containing the sequence gt in white represent the splice donors of *+R1* (5′-most gt) and *tR1* (3′-most) to the second exon of *Dnahc8*, containing the first and second in-frame methionines (not shown). Underlined sequences represent three exons of a testis-expressed antisense mRNA in part complementary to the *Dnahc8* 5′ UTR. Black boxes containing the sequences ac and ct represent the reverse complements of canonical splice donors and acceptors, respectively. The light blue rectangle containing the sequence tttatt represents the reverse complement of the canonical polyadenylation signal. Vertical dashed lines represent genomic (intronic) sequence not shown. (B) The 721-bp antisense cDNA, showing all of its possible ORFs (underlined with translational start and stop codons highlighted in yellow and blue, respectively). The underlined sequence highlighted in purple is the polyadenylation signal.

reverse transcription, translation initiation from MET2 (or any of several downstream in-frame AUGs) would most likely occur by a cap-independent internal ribosome entry site (*IRES*) mechanism (Sachs, 2000; Hellen and Sarnow, 2001).

Organization of DNAHC8 isoforms: The stem domain. Although our current data suggest that most of the C-terminally distinct DNAHC8 isoforms have different N termini, the position of a translation initiation site downstream from MET1 is not yet known. Therefore, in analyzing the amino acid sequence of DNAHC8, we will refer to MET1 as the N terminus of all isoforms. The + and *t* alleles of 3L- and 3S-ended mRNA isoforms of *Dnahc8* would thus contain ORFs of 4731 and 4202 amino acids (not including stop codons), respectively. The putative isozyme with the long C terminus and a short N terminus would be 4565 amino acids, based on the third in-frame methionine (MET3) being its presumed translational start site (see below under description of the N-terminal extension). In addition, differences between *Dnahc8*^{t} and *Dnahc8*$^{+}$ coding regions (a total of 18 *t*-specific missense mutations) will be addressed elsewhere.

A dynein heavy chain can be subdivided into a long, flexible stem or tail domain consisting of approximately the

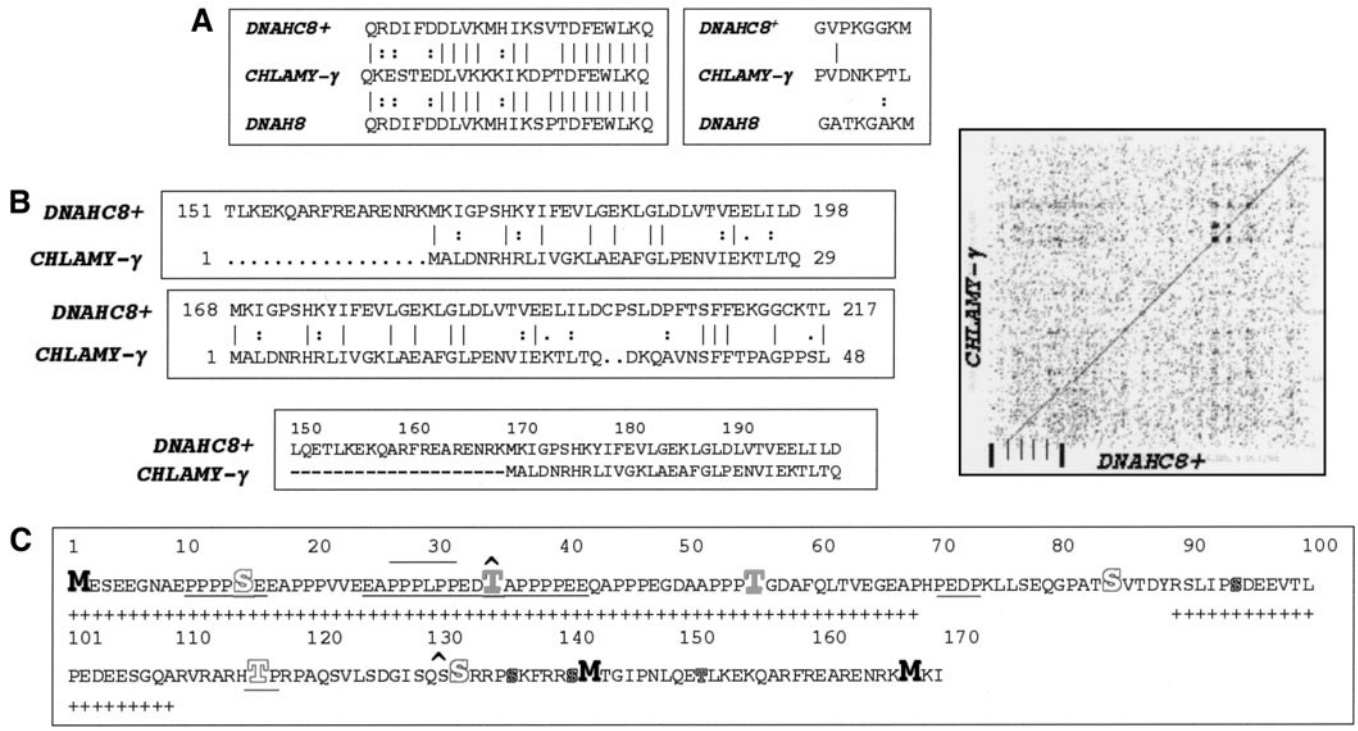

FIG. 7. Homology of the DNAHC8 and DNAH8 unique γ chain identifier sequences to the *Chlamydomonas* γ heavy chain sequence recognized by the γ chain-specific monoclonal antibody, 12γB (left), and lack of homology of the DNAHC8 and DNAH8 stem P-loops to the aligned sequence from the stem of *Chlamydomonas* γ. (B) Alignment of the N terminus of the *Chlamydomonas* γ heavy chain sequence with DNAHC8 by GAP (top left), BESTFIT (middle left), CLUSTAL W (bottom left), and COMPARE (right). All methods show that the third in-frame methionine of DNAHC8 (residue 168) aligns with the N-terminal methionine of the *Chlamydomonas* γ heavy chain. (C) The N-terminal extension of DNAHC8 with the first three in-frame methionines (tall, black Ms). +s beneath sequences represent the extent of PEST domains; outlined, small serines and threonines represent sites of probable serine/threonine phosphorylation; tall gray threonines represent potential sites of *O*-linked glycosylation; and tall, gray outlined serines and threonines represent putative sites of both *O*-linked phosphorylation and glycosylation. Vertical arrowheads above T_{34} and S_{131} represent potential sites of phosphorylation and glycosylation if the threshold for modification is reduced from >0.9 to >0.8 on a 0–1.0 scale. Underlined sequences represent sites of potential SH3 domain ligands and the horizontal lines above residues 27–30 and below residues 116–117 mark the sites of potential WW domain ligands of different types.

N-terminal 33–40% of the molecule, and a toroidally shaped motor or head unit (approximately the C-terminal 60–67%). The stem regions of dynein heavy chains are the more divergent of the two segments of these proteins (Kandl *et al.*, 1995). This presumably signifies that the stem encodes chain-type-specific functions (King, 2000b). Indeed, the stem of DNAHC8 exhibits less overall amino acid sequence similarity to other heavy chains than does the motor unit, although it is, for the most part, more closely related to the *Chlamydomonas* γ heavy chain, its human ortholog, DNAH8 (GenBank AF356519), and a *Mus* axonemal γ chain paralog, DNAH5 (GenBank AF466705), than to a *Mus* axonemal β chain or a rat cytoplasmic heavy chain (88/91% identical/similar to human DNAH8, 47/64% identical/similar to DNAH5, and 38/50% identical/similar to *Chlamydomonas* γ, but only 26/37% identical/similar to both *M. musculus* LRD, an axonemal β-type chain [GenBank NM010060] and *R. norvegicus* MAP1C, a cytoplasmic heavy chain [GenBank L08505]). In fact, the closer to the N terminus, the more divergent the sequence of DNAHC8 becomes, suggesting that the N-terminal portion of the stem may encode species- and/or cell-type-, rather than simply chain-type-specific functions.

This last point is best illustrated by analyzing different parts of the DNAHC8 stem. First, the C-terminal-most 100 amino acids of the DNAHC8 stem contain a unique γ chain identifier (residues 2003–2026), a sequence with unusually high identity/similarity (71/83%) to an epitope from the stem of the *Chlamydomonas* γ heavy chain (residues 1735–1758), previously shown to be the target of the γ chain-specific monoclonal antibody, 12γB (Fig. 7A; Wilkerson *et al.*, 1994). The presence of this highly conserved peptide in the stem of DNAHC8 suggests that some isoforms of *Dnahc8* may encode γ heavy chain-specific functions, such

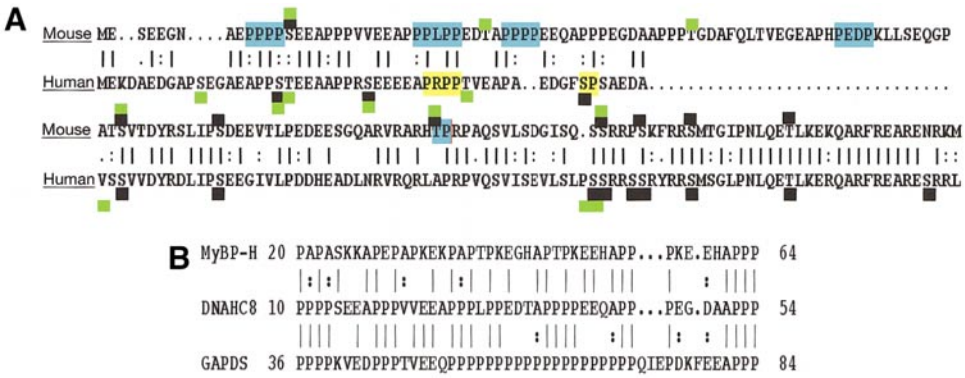

FIG. 8. (A) Alignment of the N termini of the mouse (top) and human (bottom) divergent *Dnahc8* isoforms. Sequences embedded in blue (mouse) or yellow (human) backgrounds represent potential proline-rich domain-binding ligands. Note that only one of these is fairly well conserved between mouse and human. Black rectangles represent likely sites of *O-phosphorylation*, while green rectangles represent likely sites of *O-GlcNacylation*. (B) Alignment of part of the N terminus of the divergent isoform of DNAHC8 (middle) with parts of the extended N termini of chicken Myosin Binding Protein-H (top; 53% identity/64% similarity) and the mouse sperm-specific isoform of glyceraldehyde 3-phosphate dehydrogenase (bottom; 60% identity/69% similarity).

as localization, to the axonemal dynein outer arm complex. This epitopic "ID tag" is even more highly conserved in the human orthologous protein, DNAH8 (75/88% identity/similarity to *Chlamydomonas* γ), although somewhat less well conserved in the paralogous DNAH5 (54/75% identity/similarity to *Chlamydomonas* γ).

Second, a slightly degenerate P-loop or nucleotide binding site (GVPKGGKM) is located between amino acids 237 and 244 of DNAHC8 (Fig. 7A). Unlike the P-loops in the motor unit of DNAHC8 (see below), this P-loop is not conserved in the γ heavy chain of *Chlamydomonas* nor in DNAH5, but appears to be fairly well conserved in human DNAH8 (GATKGAKM). N-terminal P-loops have been found in several (but not all) β axonemal dynein heavy chains (Kandl *et al.*, 1995; Supp *et el.*, 1999). This suggests that part of the stem of DNAHC8 might express functions other than γ-specific ones.

Third, alignment of the longest possible DNAHC8 amino acid sequence with the *Chlamydomonas* γ heavy chain sequence by either GAP, BESTFIT, COMPARE, or CLUSTAL W shows that the start codon of the *Chlamydomonas* protein is in perfect register with the third in-frame methionine of DNAHC8 at residue 168 (Fig. 7B). From that point to the two C termini, these two protein sequences exhibit good homology, displaying an overall identity/similarity of 45/56% to one another. Strikingly, the amino acid sequence consisting of residues 1–167 of DNAHC8 exhibits no significant similarity to any published heavy chain sequence except for its human ortholog with a comparable, though nonidentical, N terminus (Fig. 8A), and its γ chain paralog, DNAH5, which shows no similarity until the 128th residue of DNAHC8. This suggests that functions unique to mammalian sperm flagella may be encoded in the N-terminal-most part of the DNAHC8 stem.

A further analysis of the 167-residue N-terminal-most peptide of DNAHC8 reveals that it contains two extensive PEST sequences (residues 1–68 and 90–109; Fig. 7C; Rogers *et al.*, 1986), proline-, glutamate-, serine-, and/or threonine-rich regions often found at the termini of proteins designated for ubiquitination and proteasome-specific degradation. This two-PEST sequence architecture is maintained in the N terminus of the human ortholog (data not shown). Additional analysis of this 167-amino-acid domain with the neural network-based programs, NetPhos 2.0 (Blom *et al.*, 1999) and NetOGlyc 2.0 (Hansen *et al.*, 1998), predicts with high probability (scoring >0.9 on a scale of 0.0–1.0, where the default for modification is set at >0.5) the presence of at least eight serine/threonine phosphorylation (*O-phosphate*) sites and six O-linked glycosylation (*O-GlcNac*) sites (Fig. 7C). Outside of these six potential *O-GlcNac* sites, there are only three additional putative sites predicted in the entire DNAHC8 protein. The first five of the potential *O-GlcNac* sites are contained within the first 128 residues of the DNAHC8 N terminus, and several are conserved in DNAH8, as are numerous *O-phosphate* sites (Fig. 8A). However, none of the potential *O-GlcNac* or *O-phosphate* sites in the DNAHC8 N-terminal 167 residues is conserved in the DNAH5 N terminus (data not shown). Interestingly, the first, fourth, fifth, and sixth potential *O-GlcNac* sites are identical to the first, second, fourth, and fifth putative *O-phosphate* sites in DNAHC8 (Figs. 7C and 8A), and several of these sites are among those conserved in the human ortholog (Fig. 8A). The potential for reciprocal occupancy of sites by *O-phosphate* or *O-GlcNac* suggests that switching between *O*-linked modification states may regulate the function of this protein (Cheng and Hart, 2001).

Further, the 68-residue first PEST domain encompasses a putative WW domain-binding ligand, P-$\underline{\text{PPLP}}_{27–30}$-P (Linn *et al.*, 1997) of the formin homology 1 (FH1) type (Chan *et al.*, 1996; Bedford *et al.*, 1997), and four somewhat unusual SH3

domain-binding ligands [PPPP$_{10–13}$-SE, EAP-PPLPP$_{27–31}$-ED (two overlapping ligands), DTA-PPPP$_{36–39}$-EE], their shared odd feature being that they are flanked by negatively charged residues in the positions normally (but not always) occupied by positively charged residues in class I (+XX-PXXP) or class II (PXXPX+) SH3 domain-binding ligand motifs (Kay *et al.*, 2000). Another potential SH3 domain-binding ligand (PH-PEDP$_{70–73}$-KL) lies just distal to the first PEST domain, and a WW domain of the *Pin1*-binding type is located just distal to the second PEST domain (TP$_{116–117}$). In addition, the core proline-rich sequences of these potential adaptor domain-binding ligands are either adjacent to or very close to sites possibly modified by *O-phosphate* and/or *O-GlcNac* (Fig. 7C).

It is noteworthy that the N-terminal extension of the human ortholog contains only two potential adaptor-binding ligands, only one of which aligns with a ligand in the DNAHC8 N-terminal extension, and added potential *O-phosphorylatable* or *O-GlcNacylatable* residues appear to have replaced many of the DNAHC8 prolines in the human ortholog (Fig. 8A). This may reflect species-specific differences in the way in which the functions encoded in these N termini are regulated. Additionally, although no domain related to the first 128 residues of this N-terminal region is observed in other DHCs, its similarity, particularly between residues 10 and 54 in the first PEST sequence, to a proline-rich portion of the N-terminal extension of the mouse sperm tail-specific glyceraldehyde 3-phosphate dehydrogenase isoform (GAPDS) suggests that this region may encode a protein localization function (Fig. 8B; Welch *et al.*, 1995; Bunch *et al.*, 1998). Finally, the region containing amino acids 10–54 of the DNAHC8 N-terminal extension also shows extensive homology to the domain containing residues 20–64 of the chicken Myosin Binding Protein-H (MyBP-H) N terminus (Fig. 8B). While the function of the MyBP-H N terminus is unknown, its close relationship to a part of the N-terminal extension of DNAHC8 may reflect common ancestry.

Organization of DNAHC8 isoforms: The motor unit. The motor unit of the longer DNAHC8 isoform is comprised of six AAA subunits and a C-terminal non-AAA subunit arranged in a ring. This region of DNAHC8 is less divergent than the stem, with its first AAA domain showing the most extensive sequence conservation in other heavy chains (for the entire motor unit, 95/97% identical/similar to DNAH8, 68/79% identical/similar to DNAH5, 49/60% identical/similar to *Chlamydomonas* γ, 35/47% identical/similar to *M. musculus* LRD, and 30/42% identical/similar to *R. norvegicus* MAP1C; for the first AAA domain, 100% identical to human DNAH8, 89/96% identical/similar to DNAH5, 73/80% identical/similar to *Chlamydomonas* γ, 61/70% identical/similar to *M. musculus* LRD, and 54/64% identical/similar to *R. norvegicus* MAP1C).

As is the case with the motor units of all DHCs yet examined, the DNAHC8 motor contains four putative nucleotide binding sites or P-loops (A/GXXXXGKS/T represents the canonical P-loop motif), the most N-terminal of which corresponds to the presumptive, vanadate-sensitive (or V1 photolytic), ATP binding and hydrolytic site (Fig. 9; King, 2000a). These four P-loops are contained within the N-terminal-most four AAA domains. The first P-loop (GPAGTGKT) lies between amino acids 2087 and 2094. The sequence TETTKDM (amino acids 2094–2100) immediately following the first P-loop categorizes DNAHC8 as an axonemal, rather than a cytoplasmic heavy chain. Distal to this motif is the kinesin-like MT binding domain homology (Koonce, 1997), PGYAGRQELPENL, between amino acids 2192 and 2204 (Fig. 9) nearly completely identical to its counterpart in the γ heavy chain of *Chlamydomonas*.

The second P-loop (GPSGSGKT) is located between amino acids 2369 and 2376. The third and fourth P-loops show slight degeneracy from the canonical sequence (GEQGTAKT and GVGGSGKQ) and are located between amino acids 2697 and 2704 and 3061 and 3068, respectively (Fig. 9). Situated between the 4th and 5th AAA domains (distal to the fourth P-loop) is a structurally conserved "stalk" consisting of extended antiparallel α-helices of ~53 residues proximally (~amino acids 3354–3407) and ~63 residues distally (~amino acids 3531–3594). Atop this presumed coiled-coil "stalk" is an apparently globular domain of ~123 amino acids (beginning and ending with prolines). This noncoiled region corresponds to an apparent dynein heavy chain MT binding site (Fig. 9; Gee and Vallee, 1998).

As stated earlier, the C-terminally abbreviated DNAHC8 isoform comes to an end three amino acids past the point at which it diverges from the isoform with the long C terminus, and is missing the entire C-terminal non-AAA domain and approximately the distal half of the sixth (and most C-terminal) AAA domain relative to the longer isoform (Fig. 9; King, 2000a, Mocz and Gibbons, 2001). Thus, its C-terminal three amino acids are different from their aligned counterparts in the isoform with the long C terminus (VCN vs GGW). Moreover, although absence of the C-terminal non-AAA domain has precedent in cytoplasmic DHCs of fungi (Mocz and Gibbons, 2001), this is the first report of a naturally occurring nonfungal DHC isoform missing its C-terminal non-AAA domain as well as approximately half of its sixth AAA domain. Interestingly, the human ortholog, *DNAH8*, generates a similar truncated isoform (data not shown).

DISCUSSION

Although *Dnahc8* bifunctionality was suggested previously (Fossella *et al.*, 2000), the basis of this versatility had not yet been determined. In order to better understand the mechanisms underlying this presumptive functional flexibility, we have analyzed the molecular organization and expression of *Dnahc8* in detail. We show, first, that both the + and *t* alleles of *Dnahc8* express at least two mRNA isoforms, both more closely related to the γ DHC from the *C. reinhardtii* axonemal outer arm than to other ancestral

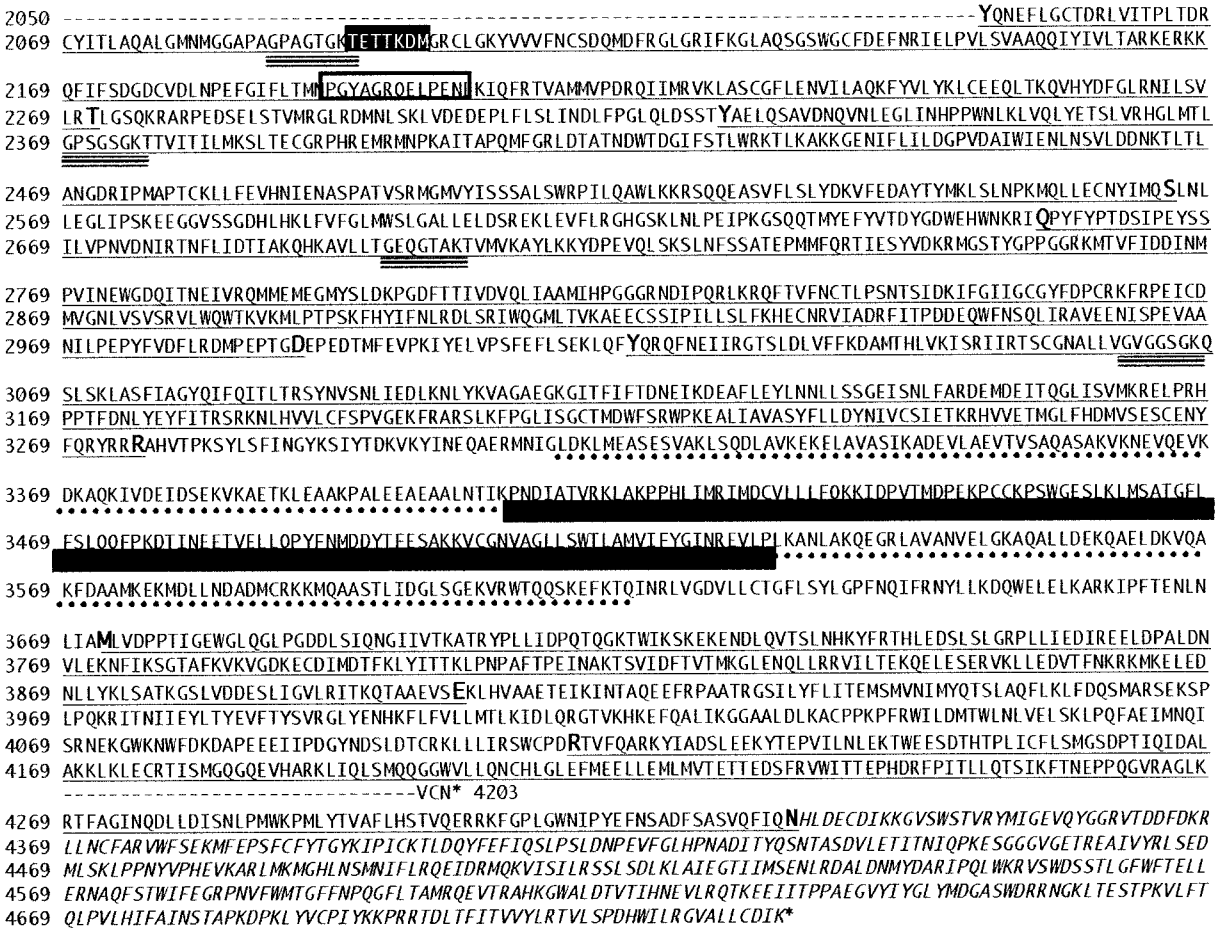

FIG. 9. Sequence of the motor units of DNAHC8 isozymes. P-loops are triple-underlined, and AAA domains are defined by single underlines flanked by bold, tall letters. White letters in a black background represent the axDHC motor unit signature sequence, and the boxed sequence corresponds to the kinesin-like microtubule binding domain homology. Dotted underlines correspond to α-helical sequences flanking the globular micotubule binding domain, represented by the black horizontal rectangle. The C-terminal non-AAA domain is italicized. Dashed line ending in "VCN* 4203" represents the C terminus of the truncated isoform of DNAHC8.

subtypes, and both highly conserved in humans. In addition, expression of both isoforms is testis-specific in the adult mouse, unlike most other previously characterized mammalian axDHCs. One of the two *Dnahc8* mRNAs is highly expressed in the testis, and appears to encode a relatively nondivergent DHC; however, the second mRNA, expressed at a rather low steady-state level, putatively encodes a DHC isoform that diverges significantly at its N and C termini (Fig. 10). Another γ axDHC gene, *Dnah5*, is also expressed in the testis (at a relatively low steady-state level), as well as in the ciliated respiratory epithelium, ependymal cells, and in nodal cilia (Olbrich *et al.*, 2002; Ibanez-Tallon *et al.*, 2002). Although *Dnah5* appears to encode a component of the axonemal outer arm in several ciliated cell types (Ibanez-Tallon *et al.*, 2002), its disruption has not yet been shown to lead to sperm flagellar dysfunction or male sterility (Olbrich *et al.*, 2002; Ibanez-Tallon *et al.*, 2002). This may be because severe *Dnah5* defects cause pronounced hydrocephaly when homozygous, and usually lead to death of the animal before it reaches sexual maturity (Ibanez-Tallon *et al.*, 2002). On the other hand, the highly expressed nondivergent isoform of *Dnahc8* may encode the predominant γ axDHC in sperm flagella, with *Dnah5* playing a minor role. Alternatively, because mammalian sperm flagella are partitioned into several compartments (middle piece, principal piece, end piece), it may be that each γ axDHC ortholog plays an equivalent role in a sperm tail compartment-specific fashion.

We have also demonstrated that both *Dnahc8* mRNA isoforms accumulate exclusively in the latter part of meiosis, indicating that translation of both must precede spermiogenesis. While we cannot rule out a role for *Dnahc8* in the late stages of male meiosis, the pattern of *Dnahc8* gene expression is typical of other testis-expressed genes whose

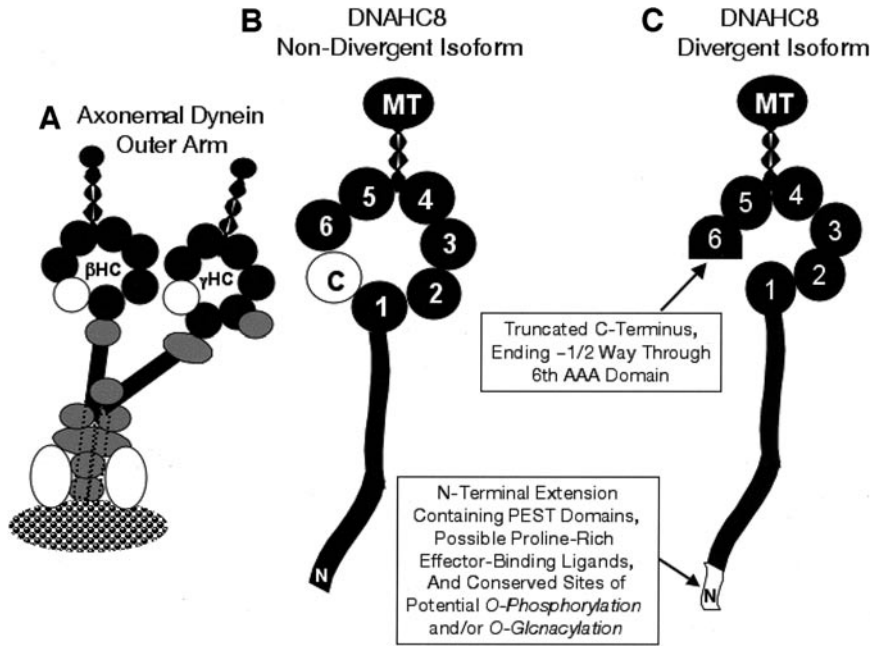

FIG. 10. Summary depictions of an axonemal dynein outer arm from a mammalian flagellum/cilium and the nondivergent and divergent axonemal dynein heavy chain isoforms of DNAHC8. (A) A typical axonemal dynein outer arm, including multiple light chains (gray), intermediate chains (white), and heterodimeric heavy chains, consisting of flexible N-terminal stems or tails (vertical gray rectangles connected to nearly vertical black rectangles, interacting with each other and at various sites with light chains), and toroidally shaped motor or head units (black circles representing 6 AAA domains and white circles representing non-AAA domain heavy chain C termini). Note that at least one light chain is thought to interact with the motor unit of the γ-heavy chain. At the base of the dynein particle is a nondynein structure (checkerboard patterned oval) anchoring the dynein to its cargo, the A-subfiber of an axonemal MT doublet. Between the fourth and fifth AAA domains of each heavy chain is a protruding coiled-coil stalk region, atop of which is a globular MT-binding site, thought to bind to and translocate along the B-subfiber of the axonemal MT doublet adjacent to the cargo MT doublet. (B, C) Cartoon representations of the DNAHC8 nondivergent and divergent isoforms. The divergent isoform contains an extended N terminus and a truncated C terminus relative to the nondivergent isoform. Although experimental evidence demonstrating the significance of these modifications is presently lacking, their conservation in human orthologous isoforms suggests that one or both modifications may play an important role in altering the stability, conformation, binding partners of the molecule, and/or ability of the molecule to translocate MTs.

protein products are active in the early part of spermiogenesis. A notable example of this expression pattern is the testis-specific linker histone, *H1t* (Steger *et al.*, 1998). Thus, *Dnahc8* is expressed at an appropriate time for one or more of its isoforms to participate in sperm tail morphogenesis, as early as its initial stages.

Although both *Dnahc8* isoforms are expressed in the latter part of meiosis, the expression pattern of each mRNA is decidedly different. The nondivergent isoform accumulates diffusely in the cytoplasm of late spermatocytes to a relatively high steady-state level, while the divergent isoform accumulates to a much lower steady-state level mainly in granules whose exact location within the cell is unclear (some of this message appears to accrue diffusely in the cytoplasm as well). In addition, its temporal pattern of expression appears to precede that of the nondivergent isoform by about one stage of the seminiferous epithelial cycle. Sequence analysis of the divergent isoform suggests that its low steady-state level may be the consequence of post-transcriptional instability induced, in part, by AU-richness and multiple AUUU sequences present in its unique 3′ UTR (Akashi *et al.*, 1991), and in part, by a testis expressed antisense mRNA whose second exon is complementary to the distinct 5′ UTR of the divergent *Dnahc8* mRNA. In the latter case, post-transcriptional gene silencing (Morel *et al.*, 2002) of *Dnahc8* would occur if the antisense mRNA is expressed in a manner that is both temporally and spatially consistent with the expression of this isoform of *Dnahc8*. However, the exact time and place of antisense expression within the testis is as yet undetermined.

The divergent *Dnahc8* mRNA isoform also encodes two extensive PEST sequences in its unique N terminus (see Fig. 10). This PEST sequence architecture is conserved in the N terminus of one of the human orthologous protein isoforms, suggesting that it may be important for regulating the stability of the encoded protein(s). Overall, this extended N terminus exhibits 0% similarity by alignment to the N termini of all other published DHCs (except for *DNAH8*, its human ortholog) for at least 128 amino acids.

Although the N termini of DHCs are normally the least conserved regions of these proteins (King, 2000b), limited similarity generally persists in this region, at least between members of a single subtype of axDHC. However, in the case of the divergent DNAHC8 isozyme, even a subtle relationship to other γ heavy chains does not appear to exist in a ≥128-residue region. Instead, part of the DNAHC8 N-terminal extension bears a more than passing resemblance to a major portion of the 103-amino-acid extended N terminus of GAPDS, a sperm tail-specific glycolytic enzyme that plays a critical role in generating energy for flagellar hyperactivation during the fertilization process. However, the extended N terminus of GAPDS has no effect on its glycolytic activity (Welch *et al.*, 1995). Instead, this "sticky finger" (Kay *et al.*, 2000) contains proline-rich ligands that apparently localize the enzyme to the principal piece of the sperm tail by anchoring it tightly (either directly or indirectly) to components of the fibrous sheath (Bunch *et al.*, 1998).

The GAPDS-homologous portion of the DNAHC8 N terminus is also replete with potential proline-rich adaptor/effector-binding ligands of several different types, suggesting that this domain is capable of interfacing, either directly or indirectly, with cytoskeletal elements. One ligand is a WW domain-binding motif of the formin homology 1 (FH1) type. Formin homologs are known to employ this ligand in the formation of complexes with several formin-binding proteins that contribute to the regulation of both actin and microtubule cytoskeletal morphogenesis and cell polarity in several cell types (Ishizaki *et al.*, 2001; Evangelista *et al.*, 2002; Sagot *et al.*, 2002).

Each of the putative proline-rich adaptor/effector-binding ligands in the DNAHC8 extended N terminus is associated with one or more nearby residues predicted to be modified by *O-phosphorylation* and/or *O-GlcNacylation*. While we lack experimental evidence for the occurrence of either *O*-linked modification at any of the residues in question in the DNAHC8 extended N terminus, it is interesting that in other systems, conditionally differential modification of a single residue by either *O-phosphate* or *O-GlcNac* has been shown to provide a role-regulating "switch," promoting or inhibiting specific molecular interactions depending on the modification state of the residue in question (Cheng and Hart, 2001). We also have not lost site of the fact that many of these sites are completely conserved in the human orthologous protein isoform. Based on this high degree of conservation as well as abundant evidence for regulation of dynein function and the binding activity of proline-rich ligands by phosphorylation (King, 2000b; Inaba *et al.*, 1998; Pelech and Sanghera, 1992), it is tempting to speculate that one or more of these conserved sites may play an important role in regulating the conformation, activity, and/or stability of the divergent DNAHC8 isoform.

Besides encoding an unusual N terminus, the divergent *Dnahc8* mRNA isoform codes for an isozyme with a truncated C terminus, due to 3′ end alternative splicing that occurs during *Dnahc8* transcription (Fig. 10). Other mammalian axonemal dynein heavy chain genes express alternatively spliced messages (Tanaka *et al.*, 1995; Supp *et al.*, 1999), but *Dnahc8* is the only DHC gene encoding an isozyme lacking more than 500 C-terminal amino acids. Similar motor unit mutations engineered into rat cytDHC genes (Gee *et al.*, 1997) have resulted in isozymes that bind MTs with increased avidity, yet fail to translocate MTs *in vitro*. The mechanism underlying these *in vitro* effects is unknown, but may be similar to one proposed for kinesin MT translocation mutants in which MT and nucleotide binding are decoupled (Song and Endow, 1998). This would suggest that the region deleted from DNAHC8 is required to catalyze the rapid transformation of basal ATPase activity into microtubule-stimulated activity necessary for the efficient release of ADP and subsequent MT translocation by dyneins. However, the *in vivo* effect of either terminal modification encoded by the divergent *Dnahc8* isoform remains enigmatic.

Despite this ongoing mystery, it is tempting to speculate that the highly divergent *Dnahc8* isoform encodes a protein that may play an early spermiogenic role in fibrous sheath (FS) morphogenesis. This hypothesis is based not only on the sequence relationship between the divergent DNAHC8 isoform N terminus and that of GAPDS, as well as the fact that a major defect observed in "whipless" sperm is severe FS dysplasia (Phillips *et al.*, 1993), but on the nature of FS development itself. The FS is a cytoskeletal structure unique to the mammalian sperm flagellar principal piece, whose presence may counteract increased internal forces in long flagella, thus helping to maintain their integrity (Fawcett, 1975; Lindemann and Kanous, 1997). It is made up of two interlocking, but independently generated and assembled components: longitudinal columns (LCs) and circumferential ribs (CRs). In order for these structures to assemble late in spermiogenesis, precursor frameworks or anlagen must first be deposited and assembled in what will become the principal piece of the developing sperm tail (Oko and Clermont, 1991). The LC precursors are assembled throughout the early part of spermiogenesis (steps 2–10) from components deposited at the distal end of the principal piece of the nascent flagellum, and assembly occurs in the retrograde direction (Oko and Clermont, 1991; Sakai *et al.*, 1986). This implies that a minus-end directed motor may be involved in the movement, distribution, and/or assembly of FSLC precursor components. Although a retrograde intraflagellar transport role for the cytDHC, DHC1b, has been established in *Chlamydomonas* flagellar assembly (Pazour *et al.*, 1999; Porter *et al.*, 1999), a role for this protein in mammalian sperm tail assembly has not been reported. If it has a role, it may be restricted to the retrograde movement of components common to the *Chlamydomonas* flagellum and other cilia and flagella. Consequently, a novel, testis-specific DHC may be necessary to coordinate the retrograde transport of flagellar components that are unique to the mammalian sperm tail. In addition, the long period of time over which FSLC framework assembly occurs (more than 1 week in the

mouse) may require a motor whose processivity is tightly regulated, perhaps by modifications to one of its heavy chains such as those encoded in the highly divergent *Dnahc8* isoform.

At first glance, the two earliest observed "whipless" defects, instability of nascent axonemal MTs and ballooning of the periaxonemal compartment (Phillips *et al.*, 1993), seem unrelated to the FS dysplasia of "whipless" sperm. However, FSLC frameworks normally connect the periaxonemal plasma membrane to outer MT doublets 3 and 8, thus providing a link between the axoneme and the plasmalemma. This raises the possibility that axonemal–plasmalemma communication is a prerequisite for maintaining the integrity of nascent axonemal MTs as well as the sleeve-like architecture of the periaxonemal compartment during early spermiogenesis. Indirect support for this proposition comes from the fact that these initial "whipless" defects are first observed at about the time that FSLC precursor assembly normally commences.

Of additional note is the fact that many of the immotile spermatozoa released from the "whipless" testis at the end of spermatogenesis contain singlet MT remnant-associated outer dense fibers (ODFs) exhibiting normal cortical and medullary ultrastructure (Pilder *et al.*, 1993; Phillips *et al.*, 1993). Since many components of the FS and the ODFs are immunologically related (Oko and Clermont, 1991), the difference in their abilities to associate with MT remnants and eventually differentiate in "whipless" sperm may reflect the fact that the precursors of each of these mammalian sperm tail structures assemble independently, and more importantly, in opposite directions, implying that their morphogenetic processes utilize different molecular motors.

ACKNOWLEDGMENTS

We thank Ms. Audrey Harmon-Smith for her technical assistance. This research was supported by NIH Grant HD-31164 (to S.H.P.).

Note added in proof. After this manuscript was accepted for publication, the following Genbank accession numbers were assigned to sequences discussed herein: AF526535 for *tL1* and AF526536 for *tL2* (see Fig. 6); AF527620 for genomic sequence around exon 1 of *Dnahc8* and exon 1 antisense DNA (see Fig. 5); AF527621 and AF527622 for two 5'UTRs of *DNAH8*, both encoding the human DNAH8 extended N-terminus; and AF527623 for an alternative 3'UTR and truncated C-terminus of human *DNAH8*.

REFERENCES

Akashi, M., Shaw, G., Gross, M., Saito, M., and Koeffler, H. P. (1991). Role of AUUU sequences in stabilization of granulocyte–macrophage colony-stimulating factor RNA in stimulated cells. *Blood* **78,** 2005–2012.

Angerer, L. M., and Angerer, R. C. (1991). Localization of mRNAs by *in situ* hybridization. *Methods Cell Biol.* **35,** 37–71.

Angerer, L. M., Cox, K. H., and Angerer, R. C. (1987). Demonstration of tissue-specific gene expression by *in situ* hybridization. *Methods Enzymol.* **152,** 649–661.

Bedford, M. T., Chan, D. C., and Leder, P. (1997). FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J.* **16,** 2376–2383.

Blom, N., Gammeltoft, S., and Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **294,** 1351–1362.

Bunch, D. O., Welch, J. E., Magyar, P. L., Eddy, E. M., and O'Brien, D. A. (1998) Glyceraldehyde 3-phosphate dehydrogenase-S protein distribution during mouse spermatogenesis. *Biol. Reprod.* **58**, 834–841.

Chan, D. C., Bedford, M. T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J.* **15,** 1045–1054.

Chemes, H. E., Brugo, S., Zanchetti, F., Carrere, C., and Lavieri, J. C. (1987). Dysplasia of the fibrous sheath: An ultrastructural defect of human spermatozoa associated with sperm immotility and primary sterility. *Fertil. Steril.* **48,** 664–669.

Cheng, X., and Hart, G. W. (2001). Alternative *O*-glycosylation/*O*-phosphorylation of serine-16 in murine estrogen receptor beta: Post-translational regulation of turnover and transactivation activity. *J. Biol. Chem.* **276,** 10570–10575.

Dutcher, S. K. (1995). Flagellar assembly in two hundred and fifty easy-to-follow steps. *Trends Genet.* **11,** 398–404.

Evangelista, M., Pruyne, D., Amberg, D. C., Boone, C., and Bretscher, A. (2002). Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast. *Nat. Cell Biol.* **4,** 32–41.

Fawcett, D. W. (1975). The mammalian spermatozoon. *Dev. Biol.* **44,** 394–436.

Fossella, J., Samant, S. A., Silver, L. M., King, S. M., Vaughan, K. T., Olds-Clarke, P., Johnson, K. A., Mikami, A., Vallee, R. B., and Pilder, S. H. (2000). An axonemal dynein at the *Hybrid Sterility 6* locus: Implications for *t* haplotype-specific male sterility and the evolution of species barriers. *Mamm. Genome* **11,** 8–15.

Gee, M. A., Heuser, J. E., and Vallee, R. B. (1997). An extended microtubule-binding structure within the dynein motor domain. *Nature* **390,** 636–639.

Gee, M., and Vallee, R. (1998). The role of the dynein stalk in cytoplasmic and flagellar motility. *Eur. Biophys. J.* **27,** 466–473.

Gibbons, I. R. (1996). The role of dynein in microtubule-based motility. *Cell. Struct. Funct.* **21,** 331–342.

Gray, N. K., and Wickens, M. (1998). Control of translation initiation in animals. *Annu. Rev. Cell Dev. Biol.* **14,** 399–458.

Hammer, M. F., Schimenti, J., and Silver, L. M. (1989). Evolution of mouse chromosome 17 and the origin of inversions associated with *t* haplotypes. *Proc. Natl. Acad. Sci. USA* **86,** 3261–3265.

Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L., and Brunak, S. (1998). NetOglyc: Prediction of mucin type *O*-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* **15,** 115–130.

Harrison, A., Olds-Clarke, P., and King, S. M. (1998) Identification of the *t* complex-encoded cytoplasmic dynein light chain *tctex1* in inner arm I1 supports the involvement of flagellar dyneins in meiotic drive. *J. Cell Biol.* **140**, 1137–1147.

Hellen, C. U., and Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **15,** 1593–1612.

Herrmann, B. G., Koschorz, B., Wertz, K., McLaughlin, K. J., and Kispert, A. (1999). A protein kinase encoded by the *t* complex

responder gene causes non-mendelian inheritance. *Nature* **402,** 141–146.

Ibanez-Tallon, I., Gorokhova, S., and Heintz, N. (2002). Loss of function of axonemal dynein *Mdnah5* causes primary ciliary dyskinesia and hydrocephalus. *Hum. Mol. Genet.* **11,** 715–721.

Inaba, K., Morisawa, S., and Morisawa, M. (1998). Proteasomes regulate the motility of salmonid fish sperm through modulation of cAMP-dependent phosphorylation of an outer arm dynein light chain. *J. Cell Sci.* **111,** 1105–1115.

Ishizaki, T., Morishima, Y., Okamoto, M., Furuyashiki, T., Kato, T., and Narumiya, S. (2001). Coordination of microtubules and the actin cytoskeleton by the Rho effector *mDia1*. *Nat. Cell Biol.* **3,** 8–14.

Kandl, K. A., Forney, J. D., and Asai, D. (1995). The dynein genes of *Paramecium tetraurelia*: The structure and expression of the ciliary beta and cytoplasmic heavy chains. *Mol. Biol. Cell* **6,** 1549–1562.

Kay, B. K., Williamson, M. P., and Sudol, M. (2000). The importance of being proline: The interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14,** 231–241.

King, S. M. (2000a). AAA domains and organization of the dynein motor unit. *J. Cell Sci.* **113,** 2521–2526.

King, S. M. (2000b). The dynein microtubule motor. *Biochim. Biophys. Acta* **1496,** 60–75.

Koonce, M. P. (1997). Identification of a microtubule-binding domain in a cytoplasmic dynein heavy chain. *J. Biol. Chem.* **272,** 19714–19718.

Lindemann, C. B., and Kanous, K. S. (1997). A model for flagellar motility. *Int. Rev. Cytol.* **173,** 1–72.

Linn, H., Ermekova, K. S., Rentschler, S., Sparks, A. B., Kay, B. K., and Sudol, M. (1997). Using molecular repertoires to identify high-affinity peptide ligands of the WW domain of human and mouse YAP. *Biol. Chem.* **378,** 531–537.

Lyon, M. F. (1984). Transmission ratio distortion in mouse *t*-haplotypes is due to multiple distorter genes acting on a responder locus. *Cell* **37,** 621–628.

Lyon, M. F. (1986). Male sterility of the mouse *t*-complex is due to homozygosity of the distorter genes. *Cell* **44,** 357–363.

Mocz, G., and Gibbons, I. R. (2001). Model for the motor component of dynein heavy chain based on homology to the AAA family of oligomeric ATPases. *Structure (Camb.)* **9,** 93–103.

Morel, J. B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* **14,** 629–639.

Oko, R. J., and Clermont, Y. (1991). Biogenesis of specialized cytoskeletal elements of rat spermatozoa. *Ann. N Y Acad. Sci.* **637,** 203–223.

Olbrich, H., Haffner, K., Kispert, A., Volkel, A., Volz, A., Sasmaz, G., Reinhardt, R., Hennig, S., Lehrach, H., Konietzko, N., Zariwala, M., Noone, P. G., Knowles, M., Mitchison, H. M., Meeks, M., Chung, E. M., Hildebrandt, F., Sudbrak, R., and Omran, H. (2002). Mutations in *DNAH5* cause primary ciliary dyskinesia and randomization of left–right asymmetry. *Nat. Genet.* **30,** 143–144.

Olds-Clarke, P., and Johnson, L. R. (1993). *t* haplotypes in the mouse compromise sperm flagellar function. *Dev. Biol.* **155,** 14–25.

Pazour, G. J., Dickert, B. L., and Witman, G. B. (1999). The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. *J. Cell Biol.* **144,** 473–481.

Pelech, S. L., and Sanghera, J. S. (1992). MAP kinases: Charting the regulatory pathways. *Science* **257,** 1355–1356.

Phillips, D. M., Pilder, S. H., Olds-Clarke, P. J., and Silver, L. M. (1993). Factors that may regulate assembly of the mammalian sperm tail deduced from a mouse *t* complex mutation. *Biol. Reprod.* **49,** 1347–1352.

Pilder, S. H., Hammer, M. F., and Silver, L. M. (1991). A novel mouse chromosome 17 hybrid sterility locus: implications for the origin of *t* haplotypes. *Genetics* **129,** 237–246.

Pilder, S. H., Decker, C. L., Islam, S., Buck, C., Cebra-Thomas, J. A., and Silver, L. M. (1992). Concerted evolution of the mouse *Tcp-10* gene family: Implications for the functional basis of *t* haplotype transmission ratio distortion. *Genomics* **12,** 35–41.

Pilder, S. H., Olds-Clarke, P., Phillips, D. M., and Silver, L. M. (1993). *Hybrid sterility-6*: A mouse *t* complex locus controlling sperm flagellar assembly and movement. *Dev. Biol.* **159,** 631–642.

Porter, M. E. (1996). Axonemal dyneins: assembly, organization, and regulation. *Curr. Opin. Cell Biol.* **8,** 10–17.

Porter, M. E., Bower, R., Knott, J. A., Byrd, P., and Dentler, W. (1999). Cytoplasmic dynein heavy chain 1b is required for flagellar assembly in Chlamydomonas. *Mol. Biol. Cell* **10,** 693–712.

Rawe, V. Y., Galaverna, G. D., Acosta, A. A., Olmedo, S. B., and Chemes, H. E. (2001). Incidence of tail structure distortions associated with dysplasia of the fibrous sheath in human spermatozoa. *Hum. Reprod.* **16,** 879–886.

Redkar, A. A., Olds-Clarke, P., Dugan, L. M., and Pilder, S. H. (1998). High-resolution mapping of sperm function defects in the *t* complex fourth inversion. *Mamm. Genome* **9,** 825–830.

Redkar, A. A., Si, Y., Twine, S. N., Pilder, S. H., and Olds-Clarke, P. (2000). Genes in the first and fourth inversions of the mouse *t* complex synergistically mediate sperm capacitation and interactions with the oocyte. *Dev. Biol.* **226,** 267–280.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science.* **234,** 364–368.

Sachs, A. B. (2000). Cell cycle-dependent translation initiation: IRES elements prevail. *Cell* **101,** 243–245.

Sagot, I., Klee, S. K., and Pellman, D. (2002). Yeast formins regulate cell polarity by controlling the assembly of actin cables. *Nat. Cell Biol.* **4,** 42–50.

Sakai, Y., Koyama, Y., Fujimoto, H., Nakamoto, T., and Yamashina, S. (1986). Immunocytochemical study on fibrous sheath formation in mouse spermiogenesis using a monoclonal antibody. *Anat. Rec.* **215,** 119–126.

Samant, S. A., Fossella, J., Silver, L. M., and Pilder, S. H. (1999). Mapping and cloning recombinant breakpoints demarcating the *hybrid sterility 6*-specific sperm tail assembly defect. *Mamm. Genome* **10,** 88–94.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46,** 659–667.

Song, H., and Endow, S. A. (1998). Decoupling of nucleotide- and microtubule-binding sites in a kinesin mutant. *Nature* **396,** 587–590.

Steger, K., Klonisch, T., Gavenis, K., Drabent, B., Doenecke, D., and Bergmann, M. (1998). Expression of mRNA and protein of nucleoproteins during human spermiogenesis. *Mol. Hum. Reprod.* **4,** 939–945.

Supp, D. M., Witte, D. P., Potter, S. S., and Brueckner, M. (1997). Mutation of an axonemal dynein affects left–right asymmetry in *inversus viscerum* mice. *Nature* **389,** 963–966.

Supp, D. M., Brueckner, M., Kuehn, M. R., Witte, D. P., Lowe, L. A., McGrath, J., Corrales, J., and Potter, S. S. (1999). Targeted deletion of the ATP binding domain of left–right dynein confirms its role in specifying development of left–right asymmetries. *Development* **126,** 5495–5504.

Tanaka, Y., Zhang, Z., and Hirokawa, N. (1995). Identification and molecular evolution of new dynein-like protein sequences in rat brain. *J. Cell Sci.* **108,** 1883–1893.

Trachtulec, Z., and Forejt, J. (2001). Synteny of orthologous genes conserved in mammals, snake, fly, nematode, and fission yeast. *Mamm. Genome* **12,** 227–231.

Welch, J. E., Brown, P. R., O'Brien, D. A., and Eddy, E. M. (1995). Genomic organization of a mouse glyceraldehyde 3-phosphate dehydrogenase gene (*Gapd-s*) expressed in post-meiotic spermatogenic cells. *Dev. Genet.* **16,** 179–189.

Wilkerson, C. G., King, S. M., and Witman, G. B. (1994). Molecular analysis of the gamma heavy chain of *Chlamydomonas* flagellar outer-arm dynein. *J. Cell Sci.* **107,** 497–506.

Wilkinson, D. G. (1995). RNA detection using non-radioactive in situ hybridization. *Curr. Opin. Biotechnol.* **6,** 20–23.

Witman, G. B. (1992). Axonemal dyneins. *Curr. Opin. Cell Biol.* **4,** 74–79.

Received for publication May 7, 2002
Revised June 27, 2002
Accepted June 28, 2002
Published online August 26, 2002