**Abstract.** We report here the cloning and sequencing of a cytoplasmic dynein heavy chain gene from the cellular slime mold *Dictyostelium discoideum*. Using a combination of approaches, we have isolated 14,318 bp of DNA sequence which contains an open-reading frame of 4,725 amino acids. The deduced molecular weight of the polypeptide predicted by this reading frame is 538,482 D. Overall, the polypeptide sequence is 51% similar and 28% identical to the recently published sequences of the β-dynein heavy chain from sea urchin flagella (Gibbons, I. R., B. H. Gibbons, G. Asai, 1988; Porter et al., 1988).

**Materials and Methods**

Unless otherwise indicated, the molecular methods were performed as described by Sambrook et al. (1989).

**cDNA Cloning**

Approximately 600,000 plaques from a *Dictyostelium* cDNA expression library (Clontech Inc., Palo Alto, CA) were screened with a polyclonal antiserum raised against the dynein heavy chain (Koonce and McIntosh, 1990). This screen identified one 4.0-kb clone that was characterized in some detail before it was used to isolate additional sequences (see Fig. 2). Restriction fragments from both the 5’ and 3’ ends of this initial clone were used to rescreen the library by hybridization, resulting in the identification of three additional clones that confirm the 5’ end of the coding sequence, and one
of sequence similarity to the published flagellar dynein sequences were the novelty of sequence at the region overlapping the previous clone, and the degree of sequence discrepancies between the PCR clones.

Double-stranded DNA sequencing was performed using the dideoxy chain termination method (Sanger et al., 1977) and Sequenase 2.0 (US Biochemicals Inc., Cleveland, OH). For most clones, nested deletions were made with exonuclease III using the Erase-a-base kit from Promega Inc. Two clones were sequenced by subcloning overlapping restriction fragments into a plasmid vector. For each PCR step, two independent clones were sequenced in their entirety. Sequence discrepancies between the PCR clones were resolved by sequencing those regions in a third independent clone, using specific primers.

Sequence assembly and initial characterization was performed using DNA Strider. Subsequent DNA and protein analysis was performed using the UWGCG sequence analysis package (Devereux et al., 1984).

**PCR Cloning**

A detailed genomic restriction map was constructed using Southern blots probed with sequence from the 3' end of the known dynein gene fragment. This map allowed us to identify enzyme sites in the 3' direction that were ~1.5 kb apart and we used these sites to asymmetrically cut and size select genomic DNA. Ligation of this DNA into a predicted orientation in a plasmid vector then gave us known DNA sequence at the 3' end with which to construct a specific downstream primer for PCR amplification.

For each PCR clone, 200 µg of genomic DNA was digested with appropriate enzymes and electrophoresed on an agarose gel. The region of the gel whose R' corresponded to the fragment of interest was excised and the DNA was glass purified. This DNA was then ligated to an equimolar amount of a vector (pGEM T7Z, Promega Corp., Madison, WI) that had been digested with the appropriate restriction endonucleases. One-fourth of this ligation reaction was added to a standard 100-µl PCR cocktail containing 20 pmol of an upstream (gene-specific) primer, and 20 pmol of a downstream (vector-specific) primer based on either the T7 or SP6 KNA polymerase binding sequences flanking the polylinker site. 35 cycles of amplification were performed using standard conditions in a thermocycler (Perkin-Elmer Corp., Norwalk, CT). The entire reaction product was electrophoresed on an agarose gel; products of the predicted size were excised, their ends were blunted with Klenow, and they were subcloned into the Smal site of pGEM. The presence of predicted internal restriction sites, the identity of sequence at the region overlapping the previous clone, and the degree of sequence similarity to the published flagellar dynein sequences were the criteria used to identify positive products.

**DNA Sequencing and Sequence Analysis**

Double-stranded DNA sequencing was performed using the dideoxy chain termination method (Sanger et al., 1977) and Sequenase 2.0 (US Biochemicals Inc., Cleveland, OH). For most clones, nested deletions were made with exonuclease III using the Erase-a-base kit from Promega Inc. Two clones were sequenced by subcloning overlapping restriction fragments into a plasmid vector. For each PCR step, two independent clones were sequenced in their entirety. Sequence discrepancies between the PCR clones were resolved by sequencing those regions in a third independent clone, using specific primers.

1. **Abbreviations used in this paper:** PCR, polymerase chain reaction.
Figure 2. Characterization of the initial cDNA (clone 1). (A) Northern blot of *Dictyostelium* total RNA probed with clone 1. (B) Southern blot of *Dictyostelium* DNA digested with the enzymes listed on top of the respective lanes. HindIII cuts once in the insert (the faint 23-kb band is due to incomplete digestion) and BgIII cuts twice. Two closely spaced EcoRV sites are present in this clone, but the 50-bp internal fragment can not be seen on this blot. (C) Immunoblots demonstrating antibody cross-reactivity. Lanes a and b show Coomassie blue-stained gel lanes of uninduced and induced cultures of *E. coli* containing the expression vector, lane b' shows an immunoblot of a sample equivalent to lane b, probed with heavy chain-purified antibodies from the original antiserum (Koonce and McIntosh, 1990). The cross-reactive bands beneath the prominent 150-kD band do not appear in blots of uninduced *E. coli*, and probably represent proteolytic fragments of the expressed polypeptide. Lane c shows a Coomassie blue-stained gel of a high-speed supernatant from *Dictyostelium*. The arrow marks the position of the dynein heavy chain. Lanes c' and c'' show immunoblot replicas of lane c probed with the preimmune (1:1,000 dilution, c') and immune serum (1:5,000 dilution, c'') from one rabbit immunized with the fusion protein. The second rabbit gave a very similar result. Lanes d and e show immunoblots with antibodies raised against the clone 1 fusion protein used to stain a blot of ATP extract from a microtubule affinity preparation of *Dictyostelium* protein. Lane d shows an unirradiated sample; lane e shows a sample that has been irradiated with UV light in the presence of vanadate and ATP. The positions of the cleavage fragments (Huv and Luv) are indicated.

Several properties of the initial clone were analyzed to determine whether it encoded a dynein-like gene product (Fig. 2). Northern blots demonstrated that this clone hybridizes to a ~14-kb RNA (Fig. 2 A), a size expected for the heavy chain message. A genomic Southern blot probed with clone 1 suggested that this is a single copy gene in *Dictyostelium* (Fig. 2 B). Clone 1 was expressed in bacteria and the resulting polypeptide was used to immunize rabbits. These antisera reacted strongly with the heavy chain from *Dictyostelium* (Fig. 2 C), but, unlike the original antiserum, they recognized epitopes found only on the lower molecular weight polypeptide defined by cleavage with UV light, ATP, and vanadate (the “LUV” polypeptide) (Gibbons et al., 1987; Koonce and McIntosh, 1990).

The derived amino acid sequence of the *Dictyostelium* dynein heavy chain gene is shown in Fig. 3. The predicted Mr of this polypeptide is 538,482 D. Like the flagellar β-dynein heavy chains, *Dictyostelium* dynein contains four GXXXXGKT/S sequences clustered in the middle third of the polypeptide. This motif represents the most conserved part of a consensus sequence for ATP-binding domains (Walker et al., 1982; Fry et al., 1986). In contrast to the flagellar dynein, however, the *Dictyostelium* sequence does not contain a fifth consensus site near its NH2-terminus. This may reflect a functional difference between these two types of dynein.

An alignment of the predicted amino acid sequences for *Dictyostelium* and sea urchin dynein is shown in Fig. 4. Using the UWGCG program BESTFIT (Devereux et al., 1984) with a gap weight of 3.0 and a length weight of 0.1, these two sequences are 51% similar and 28% identical over their entire lengths. However, as is evident in Fig. 4, this similarity is not uniformly distributed. The amino terminal 1,500-amino acids show regions of marginal similarity (19% identical), while the central region and the carboxy terminal 1,500-amino acids are 32% and 29% identical, respectively. The region of lowest sequence similarity occurs at the amino terminal 100-200 amino acids. Antibodies so far raised to the NH2-terminal region of *Dictyostelium* dynein do not cross-react with other species of cytoplasmic or axonemal dyneins, further defining its individuality (data not shown).
Figure 3. Deduced amino acid sequence of the cytoplasmic dynein heavy chain from Dictyostelium. The positions of the four putative ATP-binding consensus sequences are underlined.
The region of highest similarity between the *Dictyostelium* and sea urchin dynein sequences is in and around the putative ATP-binding consensus sequence GPAGTGKT, which shows 54% identity over 235 amino acids (* Dictyo* aa 1,924-2,159 vs. urchin aa 1,806-2,041). We have expressed a 115-amino acid peptide centered over this ATP-binding motif in *E. coli*. An antiserum raised against this peptide reacts well with dyneins from *Dictyostelium*, a mammalian cell line (HeLa), and sea urchin sperm flagella (data not shown), suggesting that this region is highly conserved among dyneins.

Fig. 5 shows direct comparisons between the two dynein sequences from the regions within and surrounding the four "segment P" consensus sequences for ATP binding (Fry et al., 1986). Gibbons et al. (1991), have suggested that the GPAGTGKT site is part of the principal hydrolytic ATP-binding site of dynein. Our data, which identify both this motif and its surroundings as regions of high sequence identity to flagellar dynein, support this assertion.

A secondary structure predicted for the *Dictyostelium* dynein sequence is presented in Fig. 6A. Given the size of the amino acid sequence, these results are hard to display, and given the inherent uncertainty in the algorithms that pre-
dict structure, they are admittedly difficult to interpret. Nonetheless, the structural information shown in Fig. 6A is consistent with a mostly globular polypeptide conformation, an observation consistent with most electron microscopic analyses of purified dynein (Warner et al., 1989), and with the structural predictions of the β-dynein heavy chain sequence (Gibbons et al., 1991; Ogawa, 1991).

The two longest regions of predicted α-helix in the Dictyostelium sequence occur between prolines 3,252–3,366 (114 amino acids) and prolines 3,491–3,630 (139 amino acids). Because the intact dynein molecule is a complex of several tightly associated polypeptides, and because coiled-coil interactions between α-helices are a fairly common motif in the formation of protein complexes, we tested whether these two regions of predicted α-helix might show diagnostic features of coiled-coil interactions (Fig. 6B and C). The “second” region of α-helix (amino acids 3,492–3,629) contains a heptapeptide repeat pattern with an enrichment of hydrophobic amino acid residues at positions a and d, as well as the periodicity of positively and negatively charged amino acids that is characteristic of a coiled-coil structural conformation (McLachlan and Karn, 1983). (Fig. 6B). The “first” region of α-helix (amino acids 3,253–3,365) shows a strong peak of hydrophobic amino acid residues at position d, but there is no second enrichment of hydrophobic residues within the 7-amino acid repeat pattern (Fig. 6C). In addition,
there is no obvious periodicity of the positively and negatively charged residues.

While these types of analyses do not rigorously demonstrate the presence or absence of a coiled–coil structural conformation, at least they provide a suggestion for a more detailed molecular characterization.

**Discussion**

Using a combination of immunological, DNA hybridization, and PCR-based techniques, we have isolated a gene sequence that encodes a dynein heavy chain from the cellular slime mold *Dictyostelium.* All the portions of this sequence so far tested hybridize to a ~14-kb RNA on Northern blots, and under high stringency conditions for Southern blotting, they recognize a single copy gene in the *Dictyostelium* genome. Bacterially expressed protein for two separate regions of this sequence have elicited antibodies in rabbits that bind with high avidity to the native dynein heavy chain. This sequence contains an open-reading frame of 4,725 amino acids, which shows substantial similarity to the flagellar β-dynein heavy chain (Gibbons et al., 1991; Ogawa, 1991).

Because part of our cloning strategy involved a multistep PCR walk along genomic DNA, we were concerned about two possible artifacts: jumping between related dynein sequences, and including sequence from introns. Three kinds of data dispel the first concern: (a) whenever adjacent clones overlapped, they contained identical nucleotide sequence; (b) all clones, except numbers 8 and 9, were used to probe Southern blots and, under high stringency conditions, all Southern blots were consistent with a single copy gene in this organism (data not shown); and (c) all the restriction sites identified by genomic Southern blotting were found in their predicted order and spacing in the final sequence. In addition, preliminary Southern blot analysis under low stringency conditions suggests that *Dictyostelium* contains only a single dynein gene (data not shown). The concern about introns can also be dismissed. In *Dictyostelium,* introns are believed to be small (~<100 bp), infrequent, highly A/T rich (95%), and bounded by conserved splice sites (Kimmel and Firtel, 1982); the clones presented here do not contain sequence with these characteristics.

Because of the small size of the *Dictyostelium* genome (~<50,000 kb; Kimmel and Firtel, 1982), the scarcity of introns in this organism, and the various levels of selection in the procedure, the PCR-based technique we have adapted was a rather quick and efficient method to walk in a given direction along a chromosome. One might expect this procedure to be a useful alternative in cloning the 5' and 3' nontranscribed regions of genes in organisms with genomes similar to *Dictyostelium.*

With the exception of the two flagellar dyneins (Gibbons et al., 1991; Ogawa, 1991), the sequence reported here shows no substantial similarity to any other proteins or nucleic acids in the GenBank, NBRF, or SwissProt databases, including detailed comparisons we have made with myosin (Warrick et al., 1986), kinesin (Yang et al., 1989), NCD (Endow et al., 1990; McDonald and Goldstein, 1990), and dynamin (Obar et al., 1990). The COOH-terminal two-thirds of the flagellar and cytoplasmic dynein heavy chains are rather similar (32% identical), suggesting that this region comprises the essential structural and mechanochemical domain of a dynein molecule. In contrast, the low degree of similarity at the amino terminus and the first one-third of the coding sequence, suggests that this region may define some of the important functional differences between cytoplasmic and axonemal dyneins, such as sites for vesicle or kinetochore binding, or for the ATP-insensitive binding to microtubules which is characteristic of the axonemal forms (Warner et al., 1989; Gibbons, 1988; Vallee and Shpetner, 1990).

Both biochemical and ultrastructural data suggest that cytoplasmic and flagellar dyneins comprise two (or three) heavy chains and various smaller molecular weight polypeptides (Warner et al., 1989; McIntosh and Porter, 1989; Vallee et al., 1988). The heavy chains of two other mechanochemical proteins myosin II and kinesin, form dimers, at least in part, through coiled–coil interactions in their rod domains (McLachlan and Karn, 1983; de Cuevas et al., 1992). Whether the dynein heavy chains associate through a similar interaction is not known. A prediction of secondary structural features in the *Dictyostelium* sequence suggests that there are only two regions of α-helix greater than 100 amino acids, both following the fourth ATP-binding consensus sequence. Both of these regions show some degree of similarity (~17% identity) to a portion of the rod domain of several myosin heavy chains and to several regions within the kinesin rod domain. Amino acid heptad repeats that are characteristic of coiled–coil structure (McLachlan and Karn, 1983) can also be found within at least one of these regions. A similar structural motif in approximately the same position has been reported for the flagellar β-dynein sequences (Gibbons et al., 1991; Ogawa, 1991). This region may reflect a conserved feature of dynein, and might be important for the tertiary and/or quaternary structure of the native molecule, either by providing some three-dimensional folding information or through association with another heavy chain or conserved accessory protein.

We are now working with this dynein sequence in an attempt to map important functional domains and to develop an understanding of dynein's role in mitosis or other cellular motile processes.

A number of people provided invaluable technical expertise. In particular, we wish to thank Tom Hays, John Tamkun, George Golumbeski, and Marc Perry for their help in answering our infinite questions. Rex Chisholm, Michiel Van Lookeren Champagne, and Richard Kessin kindly provided cDNA libraries. We also thank all the members of the McIntosh Lab, in particular Corey Nisow, Vivian Lombillo, Eugeni Vaisberg, and Bonnie Neighbors for their help and support.

This work was supported in part by grants from the National Institutes of Health (GM36663 to J. R. McIntosh) and by the Damon Runyon-Walter Winchell Cancer Research Foundation (DRG-980 to M. P. Koonce). The sequence data reported here (which represents 0.03% of the *Dictyostelium* genome) are available from EMBL/GenBank/DDJB under accession number Z15124.

Received for publication 3 September 1992 and in revised form 18 September 1992.

**References**

de Cuevas, M., T. Tao, and L. S. B. Goldstein. 1992. Evidence that the stalk of *Drosophila* kinesin heavy chain is an α-helical coiled coil. *J. Cell Biol.* 116:957–965.

Devereux, J., P. Haebeli, and O. Smithies. 1988. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acid Res.* 12:387–395.

Endow, S. A., S. Henikoff, and L. Soller-Niedziela. 1990. Mediation of meiotic and early mitotic chromosome segregation in *Drosophila* by a protein related to kinesin. *Nature (Lond.)* 345:81–83.
