The Primary Structure and Analysis of the Squid Kinesin Heavy Chain*

Kenneth S. Kosik*, Lisa D. Orcenchuk†, Bruce Schnapp‡, Hideyo Inouye‡‡, and Rachael L. Neve**

We report the cDNA sequence of the squid kinesin heavy chain and compared the predicted amino acid sequence with that of the Drosophila heavy chain as reported by Yang, J. T., Laymon, R. A., and Goldstein, L. S. B. (1989) Cell 56, 879-889. We compared the two kinesin sequences with regard to the predicted physicochemical parameters of hydrophobicity, charge, and propensities of the secondary conformations. A comparison of the sequences from the two species reveals the head, stalk, and tail domains because a reduced degree of conservation demarcates the stalk. The charge profile indicates that the head region is nearly neutral, the stalk region acidic, and the tail is basic. The Fourier transform analysis of the hydrophobic profile of the stalk shows predominant peaks at 1/3.5 and 1/2.3, which are indexed as the second and third orders of the period 7 residue. As in the Drosophila sequence, the rod domain is divided into an amino and a carboxyl subdomain by a predicted hinge region. We show that the disposition of hydrophobic residues is distinct in these two subdomains. In particular, the heptad repeat is more regular in the aminoterminal rod domain than in the carboxyl-terminal rod domain. The tail region is positively charged, a feature that is consistent with the known electrostatic interaction between kinesin heavy chain and negatively charged surfaces such as glass coveRsips and latex beads. Three monoclonal antibodies to the kinesin heavy chain have been mapped to a region within the carboxyl terminus of the stalk.

Kinesin is a soluble ATPase that drives microtubule-based motility. This protein was initially characterized in vitro by the gliding of microtubules on kinesin-coated coverslips and by the translocation of kinesin-coated plastic beads on microtubules (Vale et al., 1985a, 1985b). Recent evidence suggests that kinesin transports vesicular organelles toward the plus end of microtubules, i.e., in the anterograde direction of nerve cell axons (Schroer et al., 1988). Kinesin has now been purified from a variety of sources including squid axoplasm and optic lobes (Vale et al., 1985a), sea urchin eggs (Scholey et al., 1985), and chicken and bovine brain (Brady, 1985; Vale et al., 1985b). The molecular weight and subunit stoichiometry have been studied best in bovine brain, in which it was found that each molecule of kinesin has a mass of approximately 400 kDa, consisting of two heavy chains (135 kDa) and two light chains (60 kDa) (Bloom et al., 1988; Kuznetsov et al., 1988).

A great deal is now known about the molecular structure of kinesin. Molecular shadowing (Amos, 1987) together with antibody labeling (Hirokawa et al., 1989; Scholey et al., 1989) has shown that each molecule of kinesin has two globular heads, each contributed by one heavy chain. The two heads, which, on the basis of proteolysis studies, have ATPase activity and exhibit ATP-dependent microtubule binding (Scholey et al., 1989; Kuznetsov et al., 1989), are attached to a stalk that ends in a flattened region termed the “fan” which binds antibodies to the light chains (Hirokawa et al., 1989). Recently, the kinesin heavy chain gene from Drosophila was cloned and sequenced (Yang et al., 1989). The sequence information, in agreement with the electron microscopic data, argues that the kinesin heavy chain is a three-domain structure. The amino-terminal domain forms a globular head containing a consensus ATP-binding domain. An α-helical stalk region follows in which the two heavy chains are predicted to form a coiled coil similar to the myosin tail. Finally, there is a carboxyl-terminal domain that could correspond to the fan. A “hinge” observed by electron microscopy was described in the stalk (Hirokawa et al., 1989; Scholey et al., 1989) which could correspond to a region in the stalk sequence in which the hydrophobic periodicity breaks down (Yang et al., 1989).

Here, we report the amino acid sequence of the kinesin heavy chain from the squid Loligo pealii deduced from several cDNAs. The motivation for this work stems from the unique advantages of the squid giant axon in providing a cell-free system for reconstituting bidirectional organelle transport (Vale et al., 1985a; Schroer et al., 1988; Schnapp and Reese, 1989). At the present time, this is the only system that can be used to reconstitute bidirectional organelle transport using purified organelles and soluble factors. Cloning and sequencing of the genes for kinesin, dynamin, and other soluble factors will provide important reagents that can be used together with squid organelles to elucidate further the molecular interactions that drive directed organelle movement. In this report, we demonstrate a conserved amino acid sequence for the kinesin heavy chain between the squid and the sequence from Drosophila reported previously (Yang et al., 1989).

MATERIALS AND METHODS

Cloning and Sequencing—Optic lobes from L. pealii, obtained at the Marine Biological Laboratories in Woods Hole, were utilized to

* This work was supported in part by National Institutes of Health Grants AG06172 and AG06601 (to K. S. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by National Institutes of Health Grant NS20824 and by the American Health Assistance Fund.

‡ Supported by National Institutes of Health Grant NS20824 and by the American Health Assistance Fund.

‡‡ Supported by National Institutes of Health Grant NS20824 and by the American Health Assistance Fund.
prepare a cDNA library in λ-phage gll1 (Young and Davis, 1983a, 1983b). The library was first screened with a squid kinesin heavy chain polyclonal antibody, provided by Michael Shechter, after protein A affinity purification of the immunoglobulin fraction and preabsorption with a Y1090 *Escherichia coli* lysate. The antibody was used at 1:100 dilution, and 3H-protein A was used for detection. In order to confirm the identity of the reactive clones, the phage phase was run on sodium dodecyl sulfate-polyacrylamide gel electrophoresis and immunoblotted with the three monoclonal antibodies described above.

**Computer-assisted Analysis**—Initial entry of the nucleotide sequence, determination of reading frames, and the species alignment were done with the Genetics Computer Group Sequence Analysis software package, version 5.2 (Devereux et al., 1984). Additional analyses were conducted on a DEC 11/780 VAX with a VAX/VMS operating system in the Image Graphics Laboratory of the Children's Hospital. The programs were written in FORTRAN, and a detailed account has been published (Inouye and Kirschner, 1990). In brief, the assigned physical and chemical parameters were propensities of the α, β, β-turn conformations (Garnier et al., 1978), hydrophobicity scale (Eisenberg et al., 1984; Cornett et al., 1987), and charge density.

**RESULTS**

Antibody labeling of the expressed products from the squid optic lobe library resulted in 11 phage-purified positive clones. Eight of these clones (four of which were independent isolates based on their unique size) were also reactive with three monoclonal antibodies specific on immunoblots for squid heavy chain kinesin (Fig. 1). All of these nonreactive with the monoclonal antibodies was full-length, additional clones were obtained by PCR (Saiki et al., 1985) using a degenerate oligonucleotide corresponding to amino acids 466–474 of the *Drosophila* sequence, and using the dideoxy chain termination method of Sanger et al. (1980) with Sequenase (United States Biochemical Co.). Both strands were sequenced on an LKB macroprep apparatus. Deoxynucleotidylation of both strands of the reverse primer (Scharf et al., 1986).

The clones from this rescreen of the library were analyzed by PCR (Dorfman et al., 1989). One μl from 200 μl of the purified phage plug was used as the template. The reaction conditions were the same as above except that all 30 cycles were done with an annealing temperature of 55 °C for the remaining 30 cycles. The PCR products were analyzed by restriction mapping and Southern hybridization to confirm their physical and chemical parameters were propensities of the α, β, β-turn conformations (Garnier et al., 1978), hydrophobicity scale (Eisenberg et al., 1984; Cornett et al., 1987), and charge density.

Fig. 1. Immunoblot of the kinesin monoclonal antibody. Lane a, CG-36 on S2; lane b, CG-36 on purified kinesin; lanes c and d, CG-39 on the same two fractions; lanes e and f, CG-45 on the same two fractions; lane g, purified kinesin transferred to nitrocellulose and stained with Amido Black; lane h, S2 transferred to nitrocellulose and stained with Amido Black; lane i, high molecular weight standards (Bio-Rad) are, from the top: myosin (200,000), β-galactosidase (116,250), phosphorylase B (92,500), bovine serum albumin (66,200), and ovalbumin (45,000).
Kinesin Heavy Chain Sequence

to hybridize to the clones originally identified by the polyclonal antibody but which failed to react with the monoclonal antibodies. Clone 17A from the original screen was reactive, contained the internal EcoRI site at its 3’ end, and continued beyond the initiator methionine predicted by the Drosophila sequence into a 5’-untranslated region. We were therefore able to conclude that 12A and 17A represented contiguous segments of DNA. The sequence around the initiator methionine was confirmed by an independent isolate from the library rescreen. 21 positives were reactive with the 400-bp PCR product. One clone, designated 34A, out of 10 analyzed, extended to the 5’ end of the coding region. PCR done on the phage plug using Agt11 vector sequence as the forward primer and nucleotides 184-202 as the reverse primer gave a 1.35-kilobase fragment. A clone identical to 17A would have given an 820-bp band, and therefore, 34A is an independent isolate. The sequence of this clone confirmed the AUG of 17A as the likely site of initiation of translation. The eukaryotic translation initiation signal (Kozak, 1986) CCGCCAUGG is present (Fig. 2). A 33-bp upstream region of unknown significance is highly AT rich. Although a series of adenines is present at the 3’ terminus of the clone, it is not certain that these nucleotides are the poly(A) tail since no polyadenylation signal is present.

Combining 12A and 17A, one obtains an open reading frame that extends over 3002 nucleotides (Fig. 2). The amino acid homology with Drosophila kinesin sequence is striking (Fig. 3). Comparison of the sequences by a two-dimensional plot (Fig. 4) indicates the greatest similarities in the head and tail domains. At the site that approximates the junction of the head and stalk domains (residue -400), the Drosophila sequence has inserted amino acids. Within the long stalk region, predicted to form a coiled coil α-helical structure (Yang et al., 1989), Drosophila and squid show less homology. A comparison of the physicochemical parameters is shown in Table I. The nucleotide homology is lost in the 5’- and 3’-untranslated regions.

The motif for the putative ATP-binding site, Gly-X-X-Gly-X-Gly-Lys-Thr(Ser), or a similar sequence containing Gly-Lys-Thr(Ser) (Gorbalenya et al., 1988; Hodgman, 1988; Walker et al., 1982) begins at amino acid 79. Absolute conservation is present through this sequence except in the carboxyl
Kinesin Heavy Chain Sequence

Fig. 4. Sequence comparison between squid and Drosophila kinesin heavy chain. The sequences were compared on an all-or-none basis with a window length of 7. Points were plotted when 7 out of 7 residues were identical. The vertical lines are plotted at residues 415, 561, 580, 807 and the horizontal lines at residues 437, 583, 607, and 831. The areas enclosed by those lines correspond to structural domains; head, rod I, hinge, rod II, tail.

β-sheet propensity, and the average charge becomes positive and neutral. Within this “weak spot” in 7 successive a and d positions there is only 1 hydrophobic residue. This region, which appears to correspond to a bend in the stalk by shadowing and negatively stained images, has been putatively assigned as a hinge. In this region, the hydrophobic residues face outward, and the coiled coil structure is very likely broken.

Residues 415–807 form a heptad repeat (a, b, c, d, e, f, g)n with hydrophobic residues concentrated in the a and d positions. This segment of the molecule could extend ~60 nm, consistent with the electron microscopic observations of Amos (1987), and longer than the reported length for sea urchin (Scholey et al., 1989). Fourier transform of the hydrophobic profiles for the entire stalk clearly shows their periodic disposition every 3.5 residues, which strongly suggests the presence of an α-helix that very likely forms a coiled coil fiber in the rod region. The stalk domain is subdivided into an amino-terminal rod I, a carboxyl-terminal rod II, and a hinge between them. Rods I and II contain 147 and 228 residues with p1 values of 4.4 and 5.4, respectively. Fourier transform analysis of the hydrophobicities of the individual rods shows a relatively simple intensity profile for rod I with prominent peaks of 1/3.5 and 1/2.3, whereas rod II has many peaks, most of which are not indexed by the period 7 residue. This observation suggests that there is either a longer period or that the two rods are, in fact, separate lattice domains. The coiled coil structure appears to be more uniformly built in rod I than in rod II. In support of these distinctions is the observation by Amos (1987) of some asymmetry in the diameters of the rod on either side of the hinge.

A strong positive charge at pH 7 is present at the extreme carboxyl terminus of the sequence (Fig. 5). This site of focal charge is conserved between squid and Drosophila even though the absolute amino acids in this region are not conserved. An electrostatic interaction may serve to associate this portion of the consensus sequence in which the conservative substitution of a leucine for an isoleucine occurs at position 98. A conserved lysine (position 32) is properly spaced from the ATP binding consensus sequence to replicate the trimethylated lysine in myosin, which is thought to participate in the ATPase reaction (Tong and Elzinga, 1983).

The three-domain structure of kinesin heavy chain is also demonstrated in the display of the physicochemical parameters as a function of residue number (Fig. 5). For both the amino terminus, consisting of a head region that extends from residue 1 to ~325, and the carboxyl terminus, consisting of a tail that extends from ~825 to the end, the highest propensity alternates among four conformational features. In the stalk, the highest propensity obtained is predominantly for the α-helix. The average charge at pH 7 is nearly neutral in the head domain, mostly negative in the stalk, and positive in the tail. The calculated pI values for each domain of squid kinesin are 6.4, 4.8, and 10.6, respectively; the corresponding regions in Drosophila are 6.1, 4.8, and 10.6, respectively. Inside the stalk near residue 570, the α-helix propensity drops below the...
TABLE I
Comparison of the physicochemical parameters of the kinesin heavy chains

<table>
<thead>
<tr>
<th>No.</th>
<th>Molecular mass a</th>
<th>(k)Da</th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>pl</th>
<th>Hydrophobicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-967</td>
<td>967</td>
<td>109.4</td>
<td>63.2</td>
<td>12.2</td>
<td>5.9</td>
<td>-0.19</td>
</tr>
<tr>
<td>1-414 (head)</td>
<td>414</td>
<td>45.9</td>
<td>48.3</td>
<td>21.5</td>
<td>6.4</td>
<td>-0.05</td>
</tr>
<tr>
<td>415-807 (rod)</td>
<td>393</td>
<td>45.8</td>
<td>77.9</td>
<td>2.9</td>
<td>4.8</td>
<td>-0.29</td>
</tr>
<tr>
<td>415-561 (rod I)</td>
<td>147</td>
<td>17.4</td>
<td>70.7</td>
<td>3.4</td>
<td>4.4</td>
<td>-0.31</td>
</tr>
<tr>
<td>808-807 (rod II)</td>
<td>228</td>
<td>26.7</td>
<td>86.0</td>
<td>1.3</td>
<td>5.4</td>
<td>-0.30</td>
</tr>
<tr>
<td>808-967 (tail)</td>
<td>160</td>
<td>17.8</td>
<td>64.4</td>
<td>12.5</td>
<td>10.6</td>
<td>-0.30</td>
</tr>
<tr>
<td>Drosophila*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-975</td>
<td>975</td>
<td>110.4</td>
<td>66.6</td>
<td>12.0</td>
<td>5.6</td>
<td>-0.18</td>
</tr>
<tr>
<td>1-436 (head)</td>
<td>436</td>
<td>49.4</td>
<td>57.3</td>
<td>16.7</td>
<td>6.1</td>
<td>-0.07</td>
</tr>
<tr>
<td>437-830 (rod)</td>
<td>394</td>
<td>45.8</td>
<td>83.0</td>
<td>4.1</td>
<td>4.8</td>
<td>-0.27</td>
</tr>
<tr>
<td>437-583 (rod I)</td>
<td>147</td>
<td>17.2</td>
<td>277.6</td>
<td>4.8</td>
<td>4.3</td>
<td>-0.27</td>
</tr>
<tr>
<td>607-831 (rod II)</td>
<td>225</td>
<td>26.5</td>
<td>86.7</td>
<td>4.0</td>
<td>5.5</td>
<td>-0.33</td>
</tr>
<tr>
<td>831-975 (tail)</td>
<td>145</td>
<td>16.2</td>
<td>53.1</td>
<td>15.9</td>
<td>10.6</td>
<td>-0.28</td>
</tr>
</tbody>
</table>

* Yang et al. (1989).

FIG. 5. Physical chemical parameters as a function of residue number for squid kinesin heavy chain. The parameters, from the top, are: a, charge density at pH 7 (window length, 9); b, hydrophobicity (window length, 9) (scales from Eisenberg et al., 1984); c, \(\alpha\)-helix position; d, \(\beta\)-strand position; e, \(\alpha\)-helix propensity curve (solid); \(\beta\)-strand propensity curve (dotted); f, turn position; g, coil position; h, turn propensity curve (solid); coil propensity curve (dotted) (Garmer et al., 1978). The vertical lines are plotted at residues 415 and 808 to indicate the structural domains.

FIG. 6. A, Coomassie stain of fusion proteins. Lane b is gt11 alone, and the arrow designates \(\beta\)-galactosidase; lane c is clone 14A; lane d is 7A; lane e is 12A. In each lane, the arrow designates the fusion protein. The molecular mass standards in lane a from the bottom are: lysozyme (14,300 Da), \(\beta\)-lactoglobulin (18,400 Da), carbonic anhydrase (29,000 Da), ovalbumin (43,000 Da), bovine serum albumin (68,000 Da), phosphorylase B (97,400 Da). B, lanes f-j are Western blots with antibody CG-45. Only the fusion proteins of clones 7A and 12A are recognized by the antibody.
to fit linearly in repeating sites along the wall of the microtubule. The orientation of rod I to the microtubule has been described somewhat discrepantly by Amos (1987) and Hirokawa et al. (1989). In the former case, rod I is aligned along the microtubule, whereas in the latter case, rod I is directed toward the kinesin-coated beads at some variable angle to the microtubule wall. Although our analysis concerning the 1/38 period would support the Amos model, it is also notable that the rod I, as a firmer coiled coil, might better serve as a source of tension in the bridging of two elements.

Comparisons with myosin are revealing. The graph-matrix comparison of conserved sequences among the myosins is also able to display the major divisions of the molecule. Myosin has a strong 28-residue period (McLachlan and Karn, 1983) that is not as readily observed within the array of hydrophobic intensity peaks in kinesin (Fig. 7). It was suggested that the regular 28-residue pattern of charges in the myosin rod allows the microtubule, whereas in the latter case, rod I is directed toward the kinesin-coated beads at some variable angle to the microtubule wall. Although no evidence for kinesin phosphorylation is yet available, there are several kinase canonical sequences in the molecule which are conserved. Arg<sup>2</sup> to Ser<sup>2</sup>, Arg<sup>7</sup> to Ser<sup>7</sup>, and Arg<sup>22</sup> to Ser<sup>22</sup> are canonical for type II calcium/calmodulin kinase (Pearson et al., 1985). This is correctly positioned within a cyclic AMP-dependent kinase consensus sequence. Of unknown significance is the occurrence three times of a series of 3 of the basic amino acids lysine and arginine followed by a tyrosine.

Acknowledgments—We thank L. Goldstein for making the Drosophila kinesin sequence available to us before its publication. We are grateful to T. Reese for his support of the monoclonal antibody facility where some of the antibodies were raised. Cindi Garbus-Gerach and Ruth Bechtold provided expert technical assistance.

REFERENCES
The primary structure and analysis of the squid kinesin heavy chain.
K S Kosik, L D Orecchio, B Schnapp, H Inouye and R L Neve


Access the most updated version of this article at [http://www.jbc.org/content/265/6/3278](http://www.jbc.org/content/265/6/3278)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/6/3278.full.html#ref-list-1](http://www.jbc.org/content/265/6/3278.full.html#ref-list-1)