INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
MOLECULAR CLONING AND EXPRESSION OF THE SEA URCHIN DYNEIN BETA-HEAVY CHAIN

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN BIOMEDICAL SCIENCES (BIOCHEMISTRY) DECEMBER 1995

By
Hening Ren

Dissertation Committee:
Ian R. Gibbons, Chairperson
Frederick C. Greenwood
Marguerite Volini
David Jameson
Heinz Gert de Couet
ACKNOWLEDGEMENTS

I would like to give my sincere thanks to many people who have made my graduate tenure enjoyable. My first thanks goes to the secretaries in the Department of Biochemistry and Biophysics and Kewalo lab: June Ko, Frances Okimoto and Geri Mitomi who, with their professionalism, have navigated me through the sea of paperwork and a lot of other helps. I would like to thank Rich Chock and Steve Sakata for the work they did to ensure the smooth running of experiments all year round. To the members in the Gibbons lab, I would like to give my deepest gratitude for their tutelage, friendship and help, in particular: Dr. Grace Tang for her extensive knowledge on the biochemistry and laboratory techniques of dynein, and her advice; Dr. Barbara H. Gibbons for her intellectual insight, and teaching on several experiments; Cheryl Phillipson for her immense help on laboratory techniques and supplies; Dr. Garbor Mocz for many helpful discussions and analysis of CD data. My special thanks also goes to Dr. Grace Tang and Cheryl Phillipson for their critical reading of this dissertation.

I am deeply in debt to Dr. Ian Gibbons who has guided me, encouraged me and financially supported my research in the past six years. He has also greatly improved my English writing. My sincere thanks goes to Drs. Margueate Volini, David Jamenson, Frederick Greenwood and H. Gert deCouet for teaching and serving on my committee. I should also thank Dr. Eric Rosenthal, who served in my committee until his departure to California, for much help and advice.
Finally, I would like to thank my parents for their dedication towards my education; Hexuan Ren, my brother, for his support; and Qin Zhou, my wife, for her understanding and tremendous support.
ABSTRACT

The DNA sequence that encodes sea urchin dynein β-heavy chain (β-HC) was cloned into a plasmid vector. Plasmid pD1.3 contains a single 13,168 bp cDNA insert that encodes 98% of the β-HC. The identity of the cDNA insert was verified by partial sequencing and restriction digestion, and was found in agreement with that predicted from the published sequence determined by a PCR based method. Polypeptides encoded by the cloned cDNA reacted positively with antibodies raised against native dynein.

Truncated heavy chains of up to 300 kDa were expressed in E. coli, and some were purified under denaturing conditions and refolded.

Refolding of one of them, designated RB512 and representing amino acid residues 2,992-3,501 in the β-HC, yielded a stable soluble protein with possible native-like conformation as judged by its circular dichroic spectrum and solubility in water. Connecting the four nucleotide binding sites in the middle of β-HC with the C-terminal portion of the protein, this region contains two hypothetical coiled-coil structures that was proposed to form the short projection on the dynein head. No specific association of RB512 with in vitro polymerized microtubules was observed, nor did it significantly block the restoration of beat frequency in KCl-extracted sea urchin sperm flagella by native dynein.

The nucleotide binding domains of dynein β-HC were identified by analyzing the homology between dynein isoforms.
The putative ATP binding/hydrolysis domain and a second NTP binding domain were expressed in *E. coli* as fusion or non-fusion proteins. Soluble thioredoxin-fusion proteins were obtained using a low level expression vector, accompanied by high level co-expression of *E. coli* chaperon 70 from another co-transformed plasmid.

The soluble thioredoxin-fusion proteins, purified by Ni-chelating chromatography, were prone to precipitate upon exposure to reduced pH, heat or after prolonged storage. Photo-cleavage of fusion protein representing the ATP binding domain was observed in medium containing millimolar concentrations of vanadate upon irradiation at 365 nm, indicating the presence of native-like conformation. It did not significantly affect the motility of reactivated sea urchin sperm, but slightly inhibited the restoration of beat frequency of KCl extracted sperm by native dynein.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................ iii
ABSTRACT ....................................................... v
LIST OF TABLES ............................................. x
LIST OF FIGURES ............................................ xi
LIST OF ABBREVIATIONS .................................... xiii
CHAPTER I. INTRODUCTION ................................. 1
CHAPTER II. MATERIALS AND METHODS ................. 16

A. Construction of gene-specific cDNA library and
   isolation of dynein cDNA clones ....................... 16
   1. Isolation of poly A’ RNA ............................. 16
   2. Construction of gene-specific cDNA library .. 18
   3. Transformation of E. coli host ................... 21
   4. DNA probe synthesis ................................ 22
   5. Screening of recombinants by colony hybridization .. 25
   6. Plasmid preparations ................................ 26
   7. DNA sequencing ...................................... 28

B. Construction of expression clones and expression
   of recombinant proteins ................................ 30
   1. Isolation of DNA fragments ......................... 30
   2. Generation of expression construct ............... 32
   3. Transformation of E. coli host ................... 33
   4. Screening and characterization of expression clones ... 33
   5. Large scale culture and expressing of fusion proteins .. 35

C. Purification of fusion protein ......................... 36
   1. Lysis of bacterial cells ............................. 36
   2. Purification of recombinant protein under
      denaturing conditions ............................. 37
   3. Purification of recombinant protein under
      native conditions ................................. 38

vii
D. Studies of the properties of the soluble recombinant protein

1. Sucrose density gradient
2. Vanadate-mediated photo-cleavage
3. Effect on the motility of reactivated sperm
4. Microtubule co-sedimentation assay
5. Circular dicroic study

CHAPTER III. RESULTS

A. Isolation and characterization of cDNA clones

1. Synthesis of cDNA and construction of cDNA library
2. Screening of the cDNA libraries and characterization of cDNA clones
3. Assembly of a near full-length cDNA clone

B. Expression and purification of recombinant proteins: the large fragments

1. Expression of recombinant protein in pETS vectors
2. Expression of recombinant protein in pET24 vectors
3. Refolding of AN1480 and P748

C. Expression and purification of NTP binding domains

1. Identification of the NTP binding domains
2. Expression of NTP binding domains in pET24 vector
3. Expression of the NTP binding domains as thioredoxin or glutathione-S-transferase fusion proteins
4. Expression of recombinant proteins in a low copy number plasmid vector

D. Properties of the soluble recombinant proteins

1. The aggregation status of the soluble fusion protein
2. Vanadate mediated cleavage
3. Effect on the motility of reactivated sea urchin sperm
4. Effect on in vitro polymerized microtubules
5. Circular Dichroism spectra
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Summary of cDNA libraries</td>
<td>64</td>
</tr>
<tr>
<td>2.</td>
<td>Summary of cDNA clones</td>
<td>65</td>
</tr>
<tr>
<td>3.</td>
<td>Summary of large dynein $\beta$-HC expression constructs</td>
<td>84</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The structure of an idealized axoneme</td>
<td>12</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Three-headed outer arm dynein molecules in the bouquet configuration</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Schematic diagram of a three-headed outer arm dynein interacting with two doublet microtubules of a flagellar axoneme</td>
<td>14</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>A linear map of sea urchin dynein β-heavy chain</td>
<td>15</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Autoradiograph of cDNAs separated on agarose gel</td>
<td>66</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Chromatography of cDNA on Sephacryl H-500 column</td>
<td>67</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>EcoRI restriction digestion analysis of plasmids</td>
<td>68</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>ApaI/NotI restriction digestion analysis of plasmids</td>
<td>70</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Schematic diagram for the assembly of pD1.3</td>
<td>72</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Restriction digestion analysis of pD1.3</td>
<td>74</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Analysis of expression clone el2</td>
<td>75</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Analysis of ER1470/BL21(DE3) low temperature induction</td>
<td>77</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Analysis of pER1470/pGroESL/BL21(DE3) induction</td>
<td>79</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>Western blot analysis of RB512</td>
<td>81</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>Schematic diagram for the expression of large sections of the β-heavy chain</td>
<td>83</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>Induction of large dynein expression clones</td>
<td>85</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>A typical purification of AN1480</td>
<td>87</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>The pedigree of the P-loop expression clones</td>
<td>89</td>
</tr>
</tbody>
</table>
Figure 3.15  Schematic presentation of the alignment of amino acid sequences surrounding the nucleotide-binding sites of dynein heavy chain ............ 90

Figure 3.16  Elution of on-column refolded 349H recombinant protein .................. 92

Figure 3.17  SDS-PAGE analysis of induced 272T culture ............ 93

Figure 3.18  SDS-PAGE analysis of induced 272GST culture ............ 94

Figure 3.19  pHR map .......................... 96

Figure 3.20  A typical purification of 349TH .................... 98

Figure 3.21  Sucrose gradient centrifugation analysis of 231TH ....................... 100

Figure 3.22  Vanadate-mediated photo-cleavage of 349TH .................. 102

Figure 3.23  Vanadate mediated photo-cleavage of 349TH under various conditions .................. 104

Figure 3.24  Effect of 349TH preparation on the motility of reactivated sea urchin sperm .................. 106

Figure 3.25  Far-ultraviolet circular dichroic spectra of RB512 ..................... 106

Figure 4.1  Multiple alignment of dynein heavy chain sequences in the P1 region .................... 124

Figure 4.2  Multiple alignment of dynein heavy chain sequences in the P2 and P3 region .................... 126

Figure 4.3  Multiple alignment of dynein heavy chain sequences in the P4 region .................... 128

Figure 4.4  Schematic diagram of the hypothetical projection domain .................... 130
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>β-ME</td>
<td>β-Mercaptoethanol</td>
</tr>
<tr>
<td>Tris</td>
<td>tris (hydroxymethyl) aminomethane</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>NH₄OAc</td>
<td>ammonium acetate</td>
</tr>
<tr>
<td>LSC</td>
<td>liquid scintillation counter</td>
</tr>
<tr>
<td>O.D.₆₀₀</td>
<td>optical density at 600 nm</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>polyacrylamide gel electrophoresis in the presence of SDS</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>BCIP</td>
<td>5-bromo-4-chloro-3-indolyl phosphate</td>
</tr>
<tr>
<td>NBT</td>
<td>nitro blue tetrazolium</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The cilia and flagella of eucaryotes are highly complex structures with some common features (Goodenough and Heuser, 1985a; Gibbons, 1981). Both cilia and flagella consist of a bundle of microtubules called the axoneme, as illustrated in Figure 1.1. The axoneme is surrounded by a plasma membrane and connects with the basal body just under the main boundary of a cell. Each axoneme contains two singlet central microtubules and nine outer pairs of microtubule doublets. The central pair of microtubules and the A subfiber of the doublets are made up of 13 protofilaments, consisting of alternating $\alpha$- and $\beta$-tubulin. The B subfiber of the doublets is made up of 10 or 11 protofilaments. Each A subfiber is connected to the central microtubules by radial spokes, and to the adjacent B subfiber by highly elastic nexin links. The surface of the A subfiber is decorated by rows of dyneins at regular intervals, termed the outer dynein arm and the inner dynein arm. Mechanisms that reside in the structure of the axoneme convert the chemical energy provided by the hydrolysis of ATP to the movement of cilia and flagella, as demonstrated by the reactivation of demembranated sea urchin sperm flagella in the presence of ATP (Gibbons and Gibbons, 1972).

A major component responsible for the mechanochemical energy transduction, the dynein ATPase, was isolated about thirty years ago by Gibbons (Gibbons and Rowe, 1965). Extracting demembranated Tetrahymena cilia at low ionic
strength in the presence of EDTA solubilized the arms, central tubules and radial spokes, along with almost all of the axonemal ATPase activity. In the presence of a divalent cation, such as Mg$^{2+}$, part of the solubilized ATPase activity becoming rebound to extracted axoneme accompanied by the reparation of most of the arms (Gibbons, 1963). Methods were developed to selectively extract outer arm dyneins from demembranated axonemes in high salt medium, leaving inner arms and most other structures intact (Fay and Witman, 1977; Gibbons and Fronk, 1979). In the presence of ATP, dynein translocates microtubules in such a manner that the molecule itself moves towards the minus end of microtubule, and hence is named "minus-end directed motor" (Sale and Satir, 1977; Fox and Sale, 1987). In an axoneme, the translocation by dynein arms bound on the A-subfiber against the adjacent B-subfiber in the adjacent doublets generates shearing stress between them, which are coordinated and resisted by the radial spokes and nexin links, leading to the formation and propagation of bending waves. In axoneme whose nexin links was destroyed by a brief treatment with trypsin, addition of ATP caused a disintegration of the axoneme by gradual extrusion of the doublets (Summers and Gibbons, 1971).

Selective extraction of demembranated sea urchin sperm removed most of the outer arm dynein from the doublets, resulting in a 50% decrease in beat frequency upon reactivation with ATP (Gibbons, B.H., and Gibbons, I.R., 1973). Incubation of soluble sea urchin 21S dynein with salt extracted sperm flagella restored most of the lost outer arms,
and about 80% of the beat frequency upon reactivation with ATP (Gibbons, B.H., and Gibbons, I.R., 1976, 1979). With extracted axonemal microtubule doublets, *Tetrahymena* outer arm dynein bound to sites on both A and B subfibers, and formed stable crossbridges. Addition of MgATP induced dissociation of dynein arms attached to the B subfiber, whereas dynein arms bound to the A subfiber remained attached (Takahashi and Tonomura, 1978; Mitchell and Warner, 1980). Hence, different sites exist on the dynein that are responsible for ATP-sensitive and ATP-insensitive binding of microtubules. Dynein can also bind to and crossbridge in vitro assembled microtubules; addition of ATP induces dissociation of the crossbridges (Haimo et al., 1979; Haimo, 1982). The asymmetric nature of microtubule structure is recognized by dynein, since all the dynein arms bound to microtubules point in the same direction (Haimo and Fenton, 1984).

Examination of the soluble axonemal dyneins by electron microscopy revealed a characteristic bouquet-like appearance with two or three globular heads, joined by flexible stems to a common base (Johnson and Wall, 1983; Goodenough and Heuser, 1982, 1985), as illustrated in Figure 1.2. A projection or stalk extending from each of the globular heads that terminates in a small swelling can be seen in many preparations. The configuration of ciliary and flagellar dynein arms *in situ* assumes a more compact toadstool shape, which is a composite of two or three heavy chains joined by some of the intermediate chains. This toadstool shaped dynein attaches to the A subfiber by distally skewed subunits and...
attaches to the B subfiber through the composite projection. (Sale, Goodenough and Heuser, 1985). The interaction of dynein arms with the A and B subfibers is diagrammed in Figure 1.3.

The composition of axonemal dynein is highly complex. Different isoforms of polypeptides constitute the inner and outer arms. Depending on the species, each intact dynein molecule consists of either two ($\alpha, \beta$) or three ($\alpha, \beta, \gamma$) heavy polypeptide chains with molecular weights larger than 400 kDa and several chains of intermediate (70-125 kDa) and low (15-25 kDa) molecular weights. High-resolution gel electrophoresis reveals eight to ten high molecular weight (>300 kDa) axonemal protein bands, which may correspond to different isoforms of axonemal dynein heavy chains (Bell, Fronk and Gibbons, 1979; Piperno and Luck, 1979). More recent study of messenger RNA expression in deciliated sea urchin embryos identified 14 different genes that encode dynein heavy chains, and at least 11 of these genes participate in the regeneration of external cilia (Gibbons, B.H. et al., 1994). Seven Drosophila dynein heavy chain genes were also identified in a parallel study (Rasmusson et al. 1994).

The ATPase activity of dynein in some preparations shows a certain degree of latency, notably the 30 S dynein from Tetrahymena and 21S dynein from Tripneustes, is activated upon binding to microtubules (Blum and Hayes, 1974; Gibbons and Fronk, 1979). Exposure to organic solvents, 0.1% Triton X-100, or p-chloromercuribenzene sulfonate (PCMS) also activates
the ATPase activity, often accompanied by the dissociation of large dynein particles into smaller 10 to 14 S particles (Evans and Gibbons, 1986; Gibbons and Fronk, 1979). However, sea urchin dynein activated by Triton X-100 or PCMS completely loses its ability to recombine with and to restore the beat frequency of KCl-extracted sea urchin sperm (Gibbons and Gibbons, 1979).

Photoaffinity labelling studies with [α-³²P]8-N₃ATP on the outer arm dynein from sea urchin and *Chlamydomonas* indicated that the ATPase activity is mainly associated with the large subunits, or heavy chains, of the enzyme (Lee-Eiford, et al. 1986; Pfister, et al. 1984, 1985). Limited tryptic digestion of dynein in low salt medium cleaves the β-heavy chain at two distinct sites, and generates two defined particles, termed Fragment A and Fragment B (Figure 1.4), which can be separated by sucrose gradient centrifugation (Ogawa and Mohri, 1975; Ow et al., 1987). Fragment A retains much of the ATPase activity of the original dynein, but loses its ability to bind to dynein-depleted flagellar microtubules. It is composed principally of two polypeptides of 195 kDa and 130 kDa that are relatively resistant to further proteolysis (Bell and Gibbons, 1982; Ow et al., 1987; Mocz et al., 1988). The globular head of dynein is probably formed by this fragment. Photoaffinity labelling of dynein with [α-³²P]8-N₃ATP followed by tryptic digestion revealed the association of ³²P label with the 195-kDa peptide, indicating that this peptide contains the principle ATP hydrolysis site in dynein. Fragment B is a 110 kDa polypeptide that has no ATPase activity.
The ATPase activity of sea urchin outer arm dynein is strongly inhibited by submicromolar concentrations of inorganic vanadate, VO$_3^-$ or VO$_4^{3-}$. The inhibition is reversed by dilution into vanadate-free medium, but not by increasing the ATP concentration (Gibbons et al., 1978). Pre-steady-state kinetic analysis indicates that vanadate forms a transition state analog by binding in an uncompetitive manner at the site normally occupied by inorganic phosphate, and the more stable [dynein-ADP-Vi]-complex acts as a dead-end kinetic block (Shimizu and Johnson, 1983). Irradiation of dynein by near-ultraviolet (UV) light in the presence of monomeric vanadate, and MgATP or MgADP results in the cleavage of the heavy chain at a specific site, termed V1, to give polypeptides of apparent molecular weights of 228 kDa and 200 kDa, accompanied by a parallel loss of ATPase activity (Lee-Eifford, Ow and Gibbons, 1986; Gibbons et al., 1987). Irradiation of sea urchin dynein in the presence of oligovanadate, and in the absence of ATP cleaves the heavy chain at another specific site, termed V2, about 100 kDa from the V1 site yielding polypeptides of apparent molecular weights of 260 kDa and 170 kDa, but does not result in loss of ATPase activity (Figure 1.4) (Tang and Gibbons, 1987). Cleavage at the V2 site is inhibited by ATP, and is supported by Mn$^{2+}$, Ni$^{2+}$, Zn$^{2+}$ and Mg$^{2+}$. The sensitivity to vanadate-mediated photolysis is a general property of dyneins from many sources. Photocleavage at the V1 site is particularly well conserved, occurring in many isoforms of dynein heavy chain besides dyneins from sea urchin sperm, including flagellar
dyneins from *Chlamydomonas*, ciliary dyneins from *Tetrahymena* (King and Witman, 1987; Gibbons, I.R. and Gibbons, 1987), and cytoplasmic dyneins from mammalian brain (Vallee, et al., 1988), nematodes (Lye et al., 1987) and sea urchin eggs (Porter et al., 1988).

The sea urchin outer arm dynein was the first to have its heavy chain amino acid sequence determined by inference from cDNA (Gibbons et al. 1991; Ogawa, 1991). To date, the complete sequences of more than eight different forms of dynein heavy chains have been determined (Koonce et al., 1992; Mikami et al., 1993; Eshel et al., 1993, Wilkerson et al., 1994; Mitchell and Brown, 1994; Li et al. 1994). Partial sequences from the catalytic domains of more than a dozen dynein isoforms have also been obtained (Gibbons et al., 1994; Rasmusson et al., 1994). The predicted sizes of the complete heavy chains are in the range of 510-540 kDa. Clear sequence homology between axonemal and cytoplasmic dynein heavy chains extends over two-third of the C-terminal portion of the polypeptides. The N-terminal portion of the amino acid sequence appears to be highly divergent between the axonemal and cytoplasmic forms, but is conserved within the same group (Koonce et al., 1992; Mikami et al. 1993). A striking common feature of the dynein heavy chain sequences are four copies of the consensus motif of nucleotide-binding sites GXXXGXK(T/S/Q), or P-loops, spaced at 35- to 40-kDa intervals (Figure 1.4). The first P-loop, or P1, is identified as the principle hydrolytic ATP binding site, and is the site of vanadate mediated photocleavage under V1 conditions. The
regions surrounding it represent the most conserved part of the molecule. The fourth P-loop in the sea urchin dynein β-heavy chain, located about 100-kDa from the first P-loop, is probably the locus of V2 cleavage. Secondary structure prediction based on the amino acid sequence of the dynein heavy chain yields an α/β-type pattern along its whole length. In agreement with a circular dichroic study (Mocz and Gibbons, 1990), no extended coiled-coil tail domain exists like those in the carboxy-terminal rod region of the other motor proteins myosin and kinesin (Gibbons et al. 1991). However, two long stretches of heptad hydrophobic repeat pattern that characterize α-helical coiled coils in sea urchin β-heavy chain are found to be a conserved feature in dynein heavy chain isoforms (Koonce et al., 1992; Mikami, et al., 1993; Wilkerson et al., 1994)

The cytoplasmic forms of the dynein ATPase have been isolated and characterized from several sources (Lye et al., 1985; Paschal and Vallee, 1987; Koonse and McIntosh, 1990). The genes that encode cytoplasmic dyneins have been identified in a wide variety of eucaryotes including yeast and mammals (Eshel et al., 1993; Koonse et al., 1992; Mikami et al., 1993; Li et al., 1994; Gibbons et al., 1994). The existence of cytoplasmic dynein is nearly ubiquitous, and it appears there is only one copy of cytoplasmic dynein heavy chain gene in most organisms. This suggests that cytoplasmic dynein is more ancient than the axonemal forms. Cytoplasmic dyneins have been implicated in various forms of intracellular motility, including retrograde organelle transport (Paschal and Vallee,
1987; Shroer and Sheetz, 1991; Schnapp and Reese, 1989), endocytosis (Goltz et al., 1992), and organization of the Golgi (Corthesy-Theulaz, et al., 1992), as well as in the orientation and movement of the mitotic spindle (Pfaff et al., 1990; Steuer et al., 1990; Eshel et al., 1993). Cytoplasmic dynein is morphologically and enzymatically similar to axonemal dynein, though it has a number of distinctive properties (Vallee et al., 1988; Paschal et al., 1987; Shpetner et al., 1988). Compared to the ATPase activity of axonemal dynein which has a very strict preference for ATP over other nucleotides, cytoplasmic dynein hydrolyzes CTP, GTP and TTP at high rates, although only ATP supports in vitro motility. The ATPase activity of cytoplasmic dynein is strongly activated by low concentrations of microtubules, while much higher concentrations are required to activate ciliary and flagellar dyneins.

Considerable knowledge of dynein has been accumulated in the past thirty years since its first isolation, but relatively little is known of the tertiary structure of the molecule, and its relationship to biological function. At the center of the dynein mystery is its exceptionally large size that has been conserved throughout evolution. In addition, the four consensus nucleotide binding motifs in the heavy chain are also conserved, although only one of them appears to hydrolyze ATP (Gibbons et al., 1987). Myosin and kinesin, two well characterized motor proteins, execute essentially the same mechanochemical transduction function but with much
smaller structures (Toyoshima et al., 1987; Yang et al., 1990).

A very powerful molecular biological procedure that has been successfully employed in the research of myosin (McNally et al., 1988) and kinesin (Yang et al., 1990) is to clone the genes that encode them, and then to express the cloned genes in a suitable host. In this way, the domains that are responsible for certain biological functions can be manipulated by a variety of molecular biology techniques: such as deletion, point mutation, or fusion with a heterologous gene to form a chimeric protein. The results of domain manipulation can then be studied in a clearly defined system, such as an in vitro motility assay.

Similar techniques permitting direct access to and manipulation of putative domain structures in dynein should be of more interest than traditional biochemical and genetic techniques, for the future study of the structure-function relationship of dynein-based motility. By expressing the truncated form, molecular biology techniques could also circumvent the difficulties posed by the large size of dynein heavy chains, and reduce the complexity of polypeptide composition associated with native dynein.

Previous work in this laboratory has determined the nucleotide sequence that encodes the sea urchin dynein β-heavy chain by sequential PCR extension of a cDNA clone in a random-primed cDNA library in the bacteriophage λgt11 (Gibbons et al., 1991b). However, no single continuous cDNA clone that was suitable for expression studies was isolated. Cloning of
this cDNA would greatly simplify the procedures involving the expression of sea urchin dynein \(\beta\)-heavy chain. The purposes of this investigation are: (1) to clone the full-length cDNA sequence encoding the sea urchin dynein \(\beta\)-heavy chain into a vector suitable for the expression of polypeptides; (2) to explore the feasibility of expressing the dynein heavy chain in a heterologous host in order to obtain the polypeptide with its structure similar to that in the native dynein; (3) to explore the domain structure in dynein by such an expression system.

The experience gained in this investigation may benefit future attempts on the cloning and expression of cDNAs that encode other large polypeptides. It will also provide a system that allows direct access and manipulation of the tertiary structure of the dynein heavy chain, and to expose some aspects of the structure-functional relationship in the energy transduction process.
Figure 1.1  The structure of a schematic axoneme. Cross-sectional view showing the major structures of a schematic axoneme. For clarity, the individual protofilaments of microtubules are indicated in only one central microtubule and one outer A subfiber-B subfiber doublet microtubule. (from Goodenough and Heuser, 1985; Darnell, Lodish and Baltimore, 1986)
Figure 1.2  Three-headed outer arm dynein molecules in the bouquet configuration

Electron micrograph of the three-headed outer arm dynein molecules from the pf-23 strain of *Chlamydomonas* in the bouquet configuration. Arrowheads point to projections from the 3 dynein heads. x 350,000.

(from Goodenough and Heuser, 1989)
Figure 1.3  Schematic diagram of a three-headed outer arm dynein interacting with two doublet microtubules of a flagellar axoneme. The dynein heads are consisted of heavy chains, and the base is formed by portions of the heavy chains, intermediate chains and light chains. The outer arm dynein decorate the A subfiber at a regular 24 nm interval. (modified from Goodenough and Heuser, 1984; Darnell, Lodish and Baltimore, 1990)
Figure 1.4 A linear map of sea urchin dynein β-heavy chain.

The principle tryptic sites are indicated as T1 - T4; the principle photolytic sites are indicated as V1 and V2. The position of four consensus phosphate binding loops are indicated by P1 - P4. Numbers above the middle line give the masses of the tryptic peptides ($Mr \times 10^{-3}$) estimated from their electrophoretic migration. Thick regions of the map indicate peptides that are resistant to tryptic digestion at 37°C. (modified from Gibbons, et al., 1991a)
A. CONSTRUCTION OF GENE-SPECIFIC cDNA LIBRARY AND ISOLATION OF DYNEIN cDNA CLONES

1. Isolation of poly A+ RNA

Partially purified messenger RNA from deciliated embryos of the sea urchin *Tripneustes gratilla* was kindly provided by Dr. Barbara H. Gibbons. The RNA was isolated by acid guanidinium thiocyanate-phenol-chloroform extraction (Chomczynski and Sacchi, 1987), and selected once for poly A+ RNA with an oligo(dT)-cellulose spun column (Pharmacia) according to the manufacturer’s recommendation. Approximately 50% of the RNA recovered from one pass of this procedure was poly(A)+ RNA. It was sufficiently pure for some applications, such as reverse transcriptase directed-polymerase chain reaction (RT-PCR). For cDNA cloning, a second round of purification was performed to boost the proportion of poly(A)+ RNA to over 90%.

A scaled down version of the Pharmacia procedure was employed to further purify the RNA for cDNA library construction. Briefly, about 250 μl of the oligo(dT)-cellulose matrix was transferred to a 0.5 ml microtube. The tube was punctured at the bottom with a sterile 14 gauge needle and placed in a 1.5 ml microtube. The matrix was pre-equilibrated by passing through 0.2 ml high-salt buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.5 M NaCl) then 0.2 ml sample buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 3 M NaCl), with
centrifugation at 350 x g for 2 minutes after each addition of buffer to remove the liquids. Twenty microliters of partially purified poly A+ RNA at 1 µg/µl were heated at 65°C for 5 minutes, and quickly chilled in ice-water. The salt concentration in the RNA sample was brought to 0.5 M in NaCl with the addition of 5 µl of sample buffer, mixed gently with the oligo(dT) matrix, incubated at room temperature for 5 minutes followed by centrifugation as above. The matrix was washed three times with 100 µl each of high-salt buffer and three times with 100 µl each of low-salt buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.1 M NaCl). The effluent collected in the 1.5 ml microtube was discarded. To elute poly(A)+ RNA, the matrix was resuspended in 40 µl RNase-free TE buffer (10 mM Tris-HCl, pH 7.8, 1 mM EDTA) buffer, prewarmed to 65°C, and centrifuged at 350 x g for 2 minutes. The elution was repeated three more times. All the fractions that contain poly A+ RNA were combined, mixed with 1 µl of glycogen solution (10 mg/ml), 16 µl of sample buffer and two volumes of isopropanol, then precipitated at -20°C overnight. Precipitates were collected by centrifugation at 16,000 x g in an Eppendorf microfuge for 10 minutes at room temperature, gently rinsed with 200 µl of 70% ethanol, microfuged for 2 minutes at room temperature, drained, then dried under vacuum. The pellet was dissolved in 5 µl TE and used for cDNA synthesis. The final yield of poly A+ RNA was estimated at 5 µg, based on the binding capacity of the oligo(dT)-cellulose (approximately 25 µg/ml bed volume) and the quantity and purity of the starting material.
2. Construction of gene-specific cDNA library

First strand cDNA synthesis

An estimated 5 μg of poly A+ RNA in 10 μl water were mixed with 2 μl dynein beta-heavy chain specific primer MPNI3 (500 μg/μl) in a 0.5 ml microtube, heated to 70°C for 10 minutes, then quickly chilled in an ice-water bath. While maintaining the tube on ice, the following reagents were added: 4.8 μl of 5x RT buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM MgCl₂), 1.0 μl of 10 mM dNTP mix (10 mM each dATP, dCTP, dGTP, dTTP), 2.0 μl of 0.1 M dithiothreitol (DTT), and 2 μl of 32p-Q'-dCTP. The mixture was warmed to 37°C for 3 minutes before 3 μl of SuperScript RT (a genetically engineered Moloney murine leukemia virus reverse transcriptase, GIBCO BRL, 200 U/μl) was added and mixed immediately by gentle vortexing for a few seconds. The reaction was incubated at 37°C for 1 hour. At the end of the incubation, 2 μl were taken and mixed with 43 μl of 20 mM EDTA and 5 μl of carrier DNA. This aliquot was labelled #1 sample and used for analysis.

Second strand synthesis

To the finished first strand reaction on ice, the following reagents were added in order: 88.2 μl of H₂O, 30 μl of 5x second strand buffer (100 mM Tris-HCl, pH 6.9, 450 mM KCl, 23 mM MgCl₂, 0.75 mM β-NAD⁺, 50 mM (NH₄)₂SO₄), 3 μl of 10 mM dNTP mix, 3 μl of E. coli DNA ligase (NEB, 3 U/μl), 3 μl of E. coli DNA polymerase I (NEB, 10 U/μl), and 1 μl of RNase H (GIBCO BRL, 2 U/μl). The ingredients were mixed gently and incubated at 16°C for 2 hours, then 4 μl of T4 DNA polymerase
(NEB, 3 U/µl) were added and the incubation continued for 10 minutes. The reaction was terminated by heating in 70°C water bath for 15 minutes to inactivate the enzymes. The aqueous phase was extracted with 50 µl water saturated phenol, followed by 200 µl CHCl₃. One hundred forty microliters of aqueous phase were retrieved, mixed thoroughly with 70 µl of 7.5 M ammonium acetate (NH₄OAc), 0.52 ml of 95% ethanol (-20°C), and centrifuged immediately at room temperature for 20 minutes at 16,000 x g. The supernatant was removed carefully. The pellet was overlaid with 0.5 ml cold 70% ethanol, centrifuged for 2 minutes at top speed, supernatant removed, air dried for 10 minutes, then dissolved in 80 µl TE.

**Methylation of cDNA**

To cDNA dissolved in 80 µl TE, the following were added: 10 µl of 1 M Tris-HCl (pH 8.0), 2 µl of 5 M NaCl, 2 µl of 0.5 M EDTA, 1 µl of 20 mM S-adenosylmethionine, 4 µl of EcoRI methylase (NEB, 40 U/µl), mixed and incubated at 37°C for one hour. The reaction was terminated by heating to 70°C for 15 minutes, extracted sequentially with 100 µl phenol and 200 µl CHCl₃. The aqueous phase was mixed with 50 µl of 7.5 M NH₄OAc, 0.37 ml 95% ethanol (-20°C), centrifuged at room temperature for 20 minutes. The pellet was rinsed with 200 µl 70% ethanol, centrifuged for 2 minutes, air dried for 10 minutes after the careful removal of the supernatant.

**EcoRI linker addition**

The methylated cDNA was dissolved in 4 µl of H₂O. To the cDNA kept on ice, the following were added in order: 4 µl of phosphorylated EcoRI linker (NEB, 1 µg/µl), 1 µl of 10x T4 DNA

19
ligase buffer (500 mM Tris-HCl, pH 7.8, 100 mM MgCl$_2$, 100 mM dithiothreitol, 10 mM ATP, 250 µg/ml bovine serum albumin), 1 µl of T4 DNA ligase (NEB, 2,000,000 units/ml), mixed gently and incubated at 16°C overnight. The ligation reaction was heated to 70°C for 15 minutes to inactivate methylase. EcoRI digestion was performed with the addition of the following reagents: 165 µl of H$_2$O, 20 µl of 10X EcoRI buffer (500 mM NaCl, 1 M Tris-HCl, pH 7.5, 100 mM MgCl$_2$, 0.25% Triton X-100, 25°C), and 5 µl of EcoRI (NEB, 20 U/µl). The digestion was incubated at 37°C for 2.5 hours then heated to 70°C for 15 minutes, extracted sequentially with 100 µl of water saturated phenol, 200 µl of CHCl$_3$. The aqueous phase was mixed with 0.5 volume 7.5 M NH$_4$OAc, 2 volumes of isopropanol, and centrifuged at 16,000 x g at room temperature for 20 minutes to precipitate cDNA. The cDNA pellet was rinsed with 200 µl 70% ethanol, centrifuged for 2 minutes, and air dried for 10 minutes after ethanol was removed. The DNA pellet was dissolved in 50 µl of TEN (TE with 100 mM NaCl), and was ready for chromatography.

**Column chromatography**

The tip of a 1 ml serology pipette was plugged with cotton and the pipette was packed with 2 milliliters of Sephacryl H-500 (Pharmacia). The column was equilibrated with TEN buffer. The cDNA sample dissolved in 50 µl TEN was applied to the top of the column, and eluted with continuous addition of TEN. Fractions of 45 to 50 µl were collected in 1.5 ml microtubes and counted in the tritium channel in a Beckman LS 7000 liquid scintillation counter (LSC) for
Cerenkov emission. DNA fractions eluted in the void volume were combined, precipitated with the addition of 0.5 volume 3 M NaAc and 2 volumes of isopropanol at -20°C for 15 minutes, centrifuged at 16,000 x g for 20 minutes to pellet DNA. The pellet was rinsed with 200 μl 70% ethanol, centrifuged for 2 minutes, and dried at 37°C for 10 minutes after the supernatant had been carefully removed. The fractionated cDNA was dissolved in 10 μl of TE.

Ligation of cDNA to the vector

Five to ten micrograms of the plasmid vector pBlueScript II SK(+) (Stratagene) were digested with EcoRI and dephosphorylated as described (Sambrook, Fritsch and Maniatis, 1989). The treated vector was precipitated with alcohol and redissolved in TE. Five microliters of cDNA were mixed with 250 ng of the treated pBluescript vector in a total volume of 17 μl, heated briefly (30 seconds) in 70°C water bath, then returned to room temperature. Two microliters of 10x T4 DNA ligase buffer and 1.0 μl of T4 DNA ligase (NEB, 400 units/μl) were added, mixed and incubated at 16°C overnight. Eighty microliters of TE buffer were then added to the ligation mixture, followed by extraction once with phenol, twice with chloroform. The aqueous phase was mixed with 5 μg of glycogen (as carrier for precipitation), 10 μl of 3M NaOAc, 2 volumes of isopropanol, and DNA was precipitated -20°C, dried, and dissolved in 16 μl of water.

3. Transformation of E.coli host

Preparation of host cells for electroporation

One liter of LB medium (1% Bacto-Tryptone, 0.5% Bacto-
yeast extract, 1% NaCl) was inoculated with single colony from E. coli strain DH5α, shaken at 37°C until the O.D.<sub>600</sub> (optical density at 600 nm) reached about 0.7. The culture was chilled on ice, and the cells were harvested by centrifugation at 5,000 x g for 10 minutes. The cells were combined, resuspended in 250 ml sterile H₂O, and collected again by centrifuging at 8,000 x g at 4°C for 15 minutes. The above wash steps were repeated twice to remove salts and other electrolytes. The cells were finally resuspended in a total volume of 2 ml of 10% glycerol, frozen in 50 μl aliquots in liquid nitrogen, and stored at -80°C.

**Transformation by electroporation**

An aliquot of frozen electroporation-competent cells was thawed and mixed with 6 μl of ligated DNA. The cells were transformed by electroporation in a 1 mm gap cuvette with the following settings in an Invitrogen electroporator: 1500 V, 75 μF, 150 Ω load resistance. Immediately after discharge, 1 ml of S.O.C medium was added to resuspend the cells. The suspension was then transferred to a 100 x 17 mm polypropylene tube, and incubated at 37°C for 1 hour with shaking. Between 200 μl to 400 μl of transformed cells were spread on 150 mm LB agar (LB medium, 1.5% agar) plate with 100 μg/ml ampicillin and incubated at 37°C overnight.

4. **DNA probe synthesis**

**Synthesis of cDNA for use as PCR template**

Typically, 2-4 μg of of partially purified poly A⁺ RNA were converted to single-stanned cDNA as describe above, using oligo (dT)₄ primer. The synthesis reaction mixture was
transferred to a Centricon-100 centrifuge concentrator (Amicon) along with 0.5 ml TE and centrifuged at 5,000 x g in an SS-34 rotor (DuPont) for 10 minutes. The centrifugation was repeated twice with the addition of 0.5 ml of TE at each time. Approximately 120 μl of retentate were collected by an inverted spin into a conical collection tube. The cDNA was stored at -20°C until used.

**cDNA-directed PCR**

Five to ten microliters of single-stranded cDNA were amplified in a 100 μl reaction with the following composition: 10 mM Tris-HCl, (pH 9.0, 25°C), 50 mM KCl, 0.05% Triton X-100, 20 pmol sense and antisense gene-specific primers, 2.5-5 units Taq DNA polymerase (Promega), and 0.5 mM dATP, dCTP, dGTP and dTTP. Thirty cycles of polymerase chain reaction (Saiki, et al. 1988; Mullis and Faloona, 1987) were performed in a Perkin-Elmer DNA Thermal Cycler 480 using the following parameters: 93°C, 1 minutes; 45°C, 1 minute; 72°C, 6 minutes. After amplification, the DNA was separated from excess primers and nucleotides by agarose gel electrophoresis.

**Preparation of activated silica**

Two hundred grams of silica (1-10 μm, Sigma) were mixed with 50 ml of concentrated hydrochloric acid and 150 ml of concentrated nitric acid and brought to 500 ml total volume. The slurry was heated to near boiling and cooled down slowly to room temperature. The liquid was decanted and the sediment was resuspended in one liter of tap water. The particles that could not be suspended were discarded. The suspension was allowed to settle for 15 minutes. Supernatant and unsettled
particles were discarded. After five washes with tap water, the silica was washed in distilled water until neutral, followed by several more washes with sterile TE buffer. At this point, most of silica particles settled within 15 minutes and the supernatant was only slightly turbid. The silica was resuspended in equal volume of TE and stored in 2 ml aliquots at 4°C, with 0.1% NaN₃ as preservative.

**Purification of DNA from agarose gel**

The original procedure of Vogelstein and Gillespie (1979) was modified to separate DNA from agarose gel. A slice of gel containing the desired DNA band was excised, mixed with 2 volumes of 8M NaI solution (8M NaI, 25 mM 2-mercaptoethanol) and heated in a 70°C water bath until the agarose was totally dissolved. One microliter of suspended silica (50% v/v) was added per microgram of DNA, and incubated on ice for 5 minutes with intermittent mixing. The silica was pelleted and washed twice in 1.0 ml ethanol-TE wash buffer (70% ethanol, 25 mM Tris-HCl, pH7.8, 5 mM EDTA) by centrifugation and resuspension. The silica pellet was then dried in the air for 10 minutes, resuspended in 25-50 μl of H₂O, and heated in a 70°C water bath for 10 minutes to elute absorbed DNA.

**DNA probe labelling**

The DNA probes were labelled with ³²P-alpha dCTP by the random primer labelling method (Feinberg and Vogelstein, 1984). A random-primed oligo-labelling kit from Pharmacia was used to label 50-200 ng of DNA probes according to the procedure recommended by the manufacturer. After the labelling reaction, excess ³²P-alpha dCTP was removed by
centrifuging through a 1 ml Sephadex G-25 spin column. Total isotope incorporation was obtained by counting two microliters of the labelled probe with 4 ml scintillation cocktail in Beckman LS 7000 liquid scintillation counter.

5. Screening of recombinants by colony hybridization

The cDNA libraries were screened by a modified colony hybridization method (Sambrook, Fritsch and Maniatis, 1989), originally described by Grunstein and Hogness (1975), with further simplification.

The agar plates with bacterial colonies developed on the surface were chilled at 4°C for a minimum of 1 hour. A nitrocellulose filter circle (0.45 μm) was overlaid onto a plate and lifted about 10 seconds later, with the colonies adhering to the surface of the filter. Several holes were punched along the edge of the filter using a 16 gauge needle for future alignment. The nitrocellulose filters were then placed on a sheet of Whatman No. 3 paper saturated with colony lysis buffer (0.3N NaOH, 0.25% SDS) for 3 minutes, colony side up, transferred to another sheet of Whatman No. 3 paper saturated with neutralizing buffer (1.0M Tris-HCl, pH 7.5, 0.5M NaCl) for 5 min, and finally dried for 30 min at 60°C oven, followed by baking under vacuum for 1 hour at 80°C to immobilize the DNA. The colony stripped plate was incubated at 37°C for 4 hours to allow the regeneration of bacterial colonies.

To prepare for hybridization, the filters were rehydrated in 2x SSC (0.3M NaCl, 0.03M sodium citrate) for 30 minutes at 60°C. Loose colony debris was wiped from the
surfaces of the filters using damp Kimwipes (Kimberlay-Clark). After the filters were washed three times in 2x SSC, they were incubated in prehybridization solution [50% formamide, 2x SSPE (0.36M NaCl, 0.02M NaH₂PO₄, 2mM EDTA), 1.5x Denhardt’s reagent, 100 μg/ml fragmented Salmon testis DNA, and 0.1% SDS] at 42°C for 30 min with gentle shaking in a rectangular plastic food container. 50 ml solution were used for up to 20 sheets of 82 mm filters or 10 sheets of 132 mm filters. ³²P labelled DNA probe was denatured by boiling for 5 min in a water bath, and added directly to the prehybridization solution. The hybridization was carried out at 42°C for 16 hours with gentle shaking. At the end of the incubation period, the filters were washed three times in 200 ml 2x SSC with 0.2% SDS at 60°C for 15 minutes each. After wiping out excessive water, the filters were sandwiched between two sheet of plastic wrap, and exposed to X-ray film overnight.

Positive clones appeared as black spots on the developed X-ray film, and the corresponding positions were carefully determined with the aid of the alignment holes made previously. An agar plug about 3 mm in diameter was taken from each positive position using the blunt end of a pasture pipette, and the bacterial cells were resuspended in 500 μl YT medium by vortexing. Additional rounds of screening were performed following the same procedure to obtain homogenous clones.

6. Plasmid preparations

Plasmid mini-preparation

One milliliter of LB medium with appropriate antibiotic
was inoculated with a single colony, and grown to saturation density with shaking at 37°C overnight. Cultivation of multiple samples were usually performed in a 24-well tissue culture plate (Falcon #3047). The culture was transferred to a 1.5 ml microtube, cells were harvested by centrifugation and resuspended in 100 µl of glucose-TE buffer (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). The cell suspension was mixed very quickly (about 2 seconds) with 200 µl of cell lysis buffer (0.2N NaOH, 2% SDS) using a vortex mixer, incubated at room temperature for 3 minutes, then 150 µl of KAc-HAc buffer (3M KAc, 6M HAc) were added, mixed immediately yet gently with the cell lysate, incubated in ice for 5 minutes, and centrifuged at 16,000 x g for 10 minutes. Four hundred microliters of clear lysate were retrieved, mixed with 0.8 ml of 8M NaI solution (8M NaI, 25 mM 2-mercaptoethanol), 10 µl of activated silica, and incubated on ice for 5 minutes with intermittent gentle shaking. The silica was pelleted by centrifugation for 1 minute at maximum speed, resuspended in 1.0 ml ethanol-TE wash buffer (70% ethanol, 25 mM Tris-HCl, pH7.8, 5 mM EDTA), then centrifuged and the liquid was aspirated. This washing step was repeated once, and the silica was dried for 10 minutes, resuspended in 50 µl of TE and heated to 70°C for 5 minutes to elute plasmid DNA.

Plasmid midi-preparation

Intermediate scale plasmid preparations were performed using a chromatography system by QIAGEN following the procedure recommended by manufacturer. Typical yield of plasmid DNA from 50-150 ml of saturated bacterial culture in
LB medium ranged from 500 to 1,100 µg. Low copy-number plasmids were amplified using chloramphenicol (50 µg/ml) or spectinomycin (50 µg/ml) as described (Sambrook, Fritsch and Maniatis, 1989).

7. DNA sequencing

Sequencing reaction

DNA sequencing was performed by the modified dideoxy chain termination method using the Sequenase 2.0 kit from United States Biochemicals (Sanger et al., 1977, Tabor and Richardson, 1987, 1989) with modifications. Non-ionic detergent was used in the reaction to prevent annealing of complementary DNA strands according to the method of Bachmann et al. (1990). In a 0.5 ml tube, 0.5 to 2 µg of double stranded DNA template was mixed with 2 µl of 5x Sequenase Buffer (200mM Tris.HCl, pH7.5, 100mM MgCl₂, 250mM NaCl), 1 µl of 5% Nonidet P-40 (NP-40, Shell Chemicals), 2 pmol sequencing primer and H₂O to give a final volume of 11 µl, boiled in a water bath for 3 minutes, then quenched in ice-water. The liquid was collected at the bottom of tube with brief centrifugation, and 6.25 µl of sequencing master mix with the following composition was added: 1 µl of 0.1 M dithiothreitol, 0.4 µl of 5x Labeling Mix (7.5µM dGTP, 7.5µM dCTP, 7.5µM dTTP), 0.6 µl H₂O, 1 µl of 5% NP-40, 0.5 µl of α-³²S-dATP, 0.5 µl Mn Buffer (0.15 M sodium isocitrate, 0.1 M MnCl₂), 2 µl of 4x Enzyme Dilution Buffer (10mM Tris.HCl, pH7.5, 5mM DTT, 0.5mg/ml BSA), 0.25 µl of Sequenase Version 2.0 enzyme (T7 DNA polymerase). The tube was gently vortexed and incubated at room temperature for 2-5 minutes, then 3.6 µl
of this labelling reaction was dispensed into 2.5 μl of Termination Mix in a 96-well plastic micro assay plate (Falcon #3911), prewarmed to 37°C on an aluminum heating block. A set of four termination reactions was prepared for ddGTP (8μM ddGTP, 50mM NaCl, 80μM each dGTP, dATP, dCTP, dTTP), ddATP (8μM ddATP, 50mM NaCl, 80μM each dGTP, dATP, dCTP, dTTP), ddCTP (8μM ddCTP, 50mM NaCl, 80μM each dGTP, dATP, dCTP, dTTP), and ddTTP (8μM ddTTP, 50mM NaCl, 80μM each dGTP, dATP, dCTP, dTTP). After termination reaction had continued for 3-5 minutes, 4 μl of Stop Solution was added and mixed thoroughly.

Gel electrophoresis

Sequencing reactions were electrophoresed on a 6% polyacrylamide gel with 8M urea (5.7% acrylamide, 0.3% bisacrylamide, 8M urea, 45 mM Tris, 45 mM boric acid, 1 mM EDTA, pH 8.5) on a GIBCO BRL Model S2 vertical sequencing gel apparatus. The anode buffer was 500 ml of 0.5 x TBE buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA, pH 8.5), and the cathode buffer was 500 ml of 0.5 x TBE buffer with 0.3 M sodium acetate. The gel was pre-electrophoresed at 60 W for 30 minutes, and 2-4 μl of sample were loaded in each well, and electrophoresed at 60 W for about 2 hours. After electrophoresis, the gel was soaked in 500 ml of gel fixer (10% methanol, 10% acetic acid) for 10 min, transferred to Whatman No.3 paper, dried under vacuum at 80°C for 30 min, and autoradiographed overnight.

Sequence analysis

Nucleotide sequence data were analyzed by the use of IntelliGenetics PC/GENE computer programs.
B. CONSTRUCTION OF EXPRESSION CLONES AND EXPRESSION OF RECOMBINANT PROTEINS

1. Isolation of DNA fragments

DNA restriction digestion

Unless indicated otherwise, restriction endonucleases were purchased from New England Biolabs (NEB) and used according to the manufacturer's instructions. Double digestion by enzymes with compatible buffers was performed by adding both enzymes at the same time. For non-compatible enzymes, the first digestion was performed with the enzyme requiring low-salt buffer, then the digestion with enzyme requiring high-salt buffer was performed in 10 times the initial volume. The digested DNA was precipitated with 2 volumes of ethanol and 0.3 M sodium acetate at -20°C for a minimum of 2 hours, and centrifuged for 10 minutes at 16,000 x g to collect the pellet. The pellet was drained, washed with 200 μl of 70% ethanol, centrifuged at 16,000 x g for 2 minutes, drained and air dried for 10 minutes.

Agarose gel electrophoresis

DNA precipitates were dissolved in 20 to 40 μl of TE, 1/10 volume of 10x gel loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF, 80% glycerol, 50 mM EDTA) was added to the sample and mixed. Samples were loaded on an agarose gel (SeaKem, FMC Bioproducts) prepared in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0), or 0.5 x TBE buffer (45 mM Tris, 45 mM borate, 1 mM EDTA, pH 8.0), or 0.5 x TTE buffer (45 mM Tris, 15 mM taurine, 1 mM EDTA, pH 8.0) and electrophoresed. Ethidium bromide was added to the gel at 0.1 μg/ml before
casting. DNA bands were visualized by illuminating with 360 nm UV light, either from a handheld unit (EN-28 lamp, Spectronics Corp, Westbury, NY) or a Fotodyne UV box (Fotodyne Inc., New Berlin, WI). Gels were photographed on Polaroid type 55 positive/negative film using a Kodak No.8 Wratten gelatin filter. Prior to further manipulation, DNA fragments were usually purified by the silica binding method as described in "DNA probe synthesis" section or by the method described below.

Purification of DNA by binding to DEAE-membrane

Recovery of DNA fragments from agarose was usually performed by binding onto anion exchanger membrane according to Dretzen et al. (1981). Intact DNA fragments of over 8 kb could be recovered with satisfactory yield. After gel separation, a DEAE-membrane (NA45, Schleicher & Schuell) strip was placed perpendicular to the side of the gel area containing the DNA band of interest. The gel was rotated 90° and the DNA was electrophoresed onto the membrane. The membrane was rinsed twice in 1 ml each TE buffer and placed in a 1.5 ml microtube. The DNA was eluted by incubating the membrane in 150 µl high salt TE buffer (TE buffer plus 2.5M NaCl) at 70°C for 15 minutes with intermittent shaking. The elution was repeated three times, the eluate was combined, diluted with 250 µl H₂O, and the DNA was precipitated with an equal volume of isopropanol at -20°C overnight. The pellet was collected by centrifugation, decanted, washed with 70% ethanol, centrifuged, decanted and dried.
2. Generation of the expression construct

**Cohesive-end ligation**

Ligation of the DNA fragment to a vector with compatible cohesive-ends was performed according to the standard method (Sambrook, Fritsch and Maniatis, 1989). Routinely, 0.5 - 2 μg of the DNA fragment was ligated with 100 ng of vector DNA, which was digested with the appropriate restriction enzyme and dephosphorylated with alkaline phosphatase in a 20 μl reaction at 16°C overnight. The ligated DNA was precipitated with sodium acetate and isopropanol, washed in 200 μl 70% ethanol, dried and dissolved in 10 μl H₂O for use in transformation by electroporation.

**Blunt-end ligation**

Typically, 1-4 μg of a phosphorylated, blunt-ended DNA fragment were combined with 200 ng of blunt-ended, dephosphorylated vector DNA, brought to 7 μl with H₂O, heated briefly in a 70°C water bath, then chilled on ice. One microliter of 10x T4 DNA ligase buffer and 1 μl of T4 DNA ligase (NEB, 400,000 units/ml) were added. Several drops of nuclease-free mineral oil were overlayed to prevent evaporation, and the tube was incubated at 16°C overnight. The ligation was diluted with 40 μl of TE, extracted with 200 μl of chloroform to separate it from the mineral oil. DNA in the aqueous phase was precipitated with sodium acetate/isopropanol at -20°C for a minimum of 2 hours, centrifuged, washed with 70% ethanol, dried and dissolved in 10 μl H₂O for use in transformation by electroporation.
3. Transformation of the E. coli host

**Transformation of expression host**

A lysogenic E. coli strain BL21(DE3), which contains a chromosomal copy of T7 RNA polymerase gene under the control of *lacUV5* promoter (Studier and Moffatt, 1986), was used as the host for protein expression when the pET series of vectors or other vectors in which T7 promoter were used. When other vector systems were used, the hosts were used according to each supplier's recommendation. The transformation was performed by electroporation, rescued in S.O.C. medium, and plated on an LB agar plate with appropriate antibiotic, and incubated at 37°C to allow colonies to develop.

**Co-transformation of plasmids**

Cells from 1.5 ml saturated culture of a bacterial clone harboring the fusion protein expression plasmid construct were harvested by centrifugation at 5,000 x g, and washed twice in 1.0 ml ice-cold sterile H2O by resuspension-centrifugation. The cells were then resuspended in 30 μl of H2O, mixed with 500 ng of a compatible plasmid that carrying the gene of bacterial chaperon proteins and electroporated. The transformants were rescued in 1.0 ml S.O.C. medium at 30°C for 1 hour, and 5 - 50 μl were spread on LB agar plates with appropriate antibiotics selecting for both plasmids.

4. Screening and characterization of expression clones

**Immunological screening of expression clones**

Colonies on LB agar plates were transferred to nitrocellulose filter circles as described in "Screening of recombinants by colony hybridization". An LB agar plate was
spread with 20 μl of 10 mM IPTG (isopropyl-β-D-thiogalactopyranoside), and the filter was placed on it, colony side up, incubated at 37°C for 2 hours to induce expression, then treated in chloroform vapor for 15 minutes, and finally incubated for a minimum of 3 hours in colony lysis buffer (100 mM Tris-HCl, pH 7.8, 150 mM NaCl, 5 mM MgCl₂, 1 μg/ml DNase I, 1 μg/ml RNase A, 40 μg/ml lysozyme, 1.5% bovine serum albumin or nonfat milk) with gentle agitation on a rocker platform to block empty binding sites on the filter. Colony debris on the surface of the filter was wiped off, and the blocked filter was washed 3 times in Tris-buffered saline (TBST buffer, 100 mM Tris-HCl, pH 7.8, 150 mM NaCl, 0.5% Tween 20). The filter was then incubated overnight with antibodies (antiserum or monoclonal antibodies) diluted with TBST plus 0.1% BSA. The filters were rinsed 3 times in TBST, incubated with alkaline phosphatase conjugated secondary antibody (anti-rabbit-AP or anti-mouse-AP, Promega corporation) at 1/7500 dilution for 1 hour, and washed 3 times in TBST. Alkaline phosphatase developer was freshly made by mixing 7.5 ml AP buffer (100 mM Tris-HCl, pH 9.0, 100 mM NaCl, 10 mM MgCl₂) with 33 μl of 50 mg/ml BCIP (5-bromo-4-chloro-3-indolyl phosphate), and 66 μl of 50 mg/ml NBT (nitro blue tetrazolium). The filters were incubated in AP developer with gentle agitation until proper contrast was obtained. The developed filter was aligned with the LB plate via the alignment holes, and positive clones were picked. An additional round of plating and screening was performed in order to obtain homogenous clones.
**Induction of small volume culture**

To each well of a 24-well tissue culture plate, 0.6 ml of LB medium with appropriate antibiotics was added, and each was inoculated with a single colony. The plate was incubated at 37°C or 30°C with gentle shaking until an O.D._600_ of 0.5-0.7 was obtained. An aliquot of 200 µl of the culture was saved as an uninduced sample. To the rest of culture, 5 µl of 10 mM IPTG was added and incubation continued for 1-2 hours.

**Analysis of induction sample by SDS-PAGE**

Protein samples were routinely analyzed by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulphate (SDS-PAGE). 200 µl of induced or uninduced culture were centrifuged at 5,000 x g in 0.5 ml micro tube for 2 minutes, and 160 µl of the supernatant were removed. The cells were resuspended, mixed with 40 µl of 5x SDS sample buffer, boiled in a water bath for 3 minutes, and centrifuged for 5 minutes at 16,000 x g. A 6% or 13 % Tris-glycine buffered SDS polyacrylamide gel (Laemmli, 1970) was cast in a Bio-Rad Mini-Protein II apparatus. About 10 µl of the cleared SDS sample were loaded on the gel, electrophoresed at 200 V for 42 minutes. The gel was then stained with Coomassie Brilliant Blue R-250 for 15 minutes, destained, and photographed on Polaroid type 55 positive/negative film using a Kodak No.22 Wratten gelatin filter.

5. **Large scale culture and expressing of recombinant proteins**

A mini-culture of the bacterial clone expressing the desired fusion protein was prepared by inoculation of a single
colony into 2 ml LB medium with appropriate antibiotics, then grown at 30°C or 37°C overnight. The cells were harvested and washed twice in 1.4 ml cold sterile water by successive steps of resuspension and centrifugation in a 1.5 ml microtube. The washed cells were resuspended and used to inoculate 0.5 to 1 liter LB medium in a 2.8-liter culture flask supplemented with antibiotics, and the culture was shaken at 30°C or 37°C.

For induction with IPTG, the cell density was monitored by measuring O.D. 600, and IPTG was added to 0.5 mM when the density of the culture reached ~0.6. The induction was then continued for 1 - 2 hour before harvesting. For clones that constitutively express the fusion protein, the culture growth was extended overnight.

C. PURIFICATION OF RECOMBINANT PROTEINS

Recombinant proteins which express a consecutive stretch of six histidine residues at the carboxyl terminus were purified by immobilized metal chelate affinity chromatography (Hochuli, 1990). The chelating adsorbant, Ni-NTA (nitrilotri-acetic acid) agarose, was purchased from QIAGEN Inc., and used according to the supplied procedure with modifications.

1. Lysis of bacterial cells

The bacterial cells were harvested by centrifugation, resuspended in 5 volumes sonication buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl or sodium acetate, 10 mM β-mercaptoethanol) with 2 mg/ml lysozyme, and incubated on ice for 30 minutes. Before the lysis of cell, protease inhibitors PMSF (phenylmethyl sulfonyle fluoride) and TAME (p-toluenesulfonyl-L-arginine methyl ester) were added at 200
ng/ml to suppress proteolytic degradation of recombinant protein. The cell suspension was sonicated intermittently at 20 watts for about 1 minute with a microprobe in a 50 ml polycarbonate centrifuge tube in an ice-water bath. The lysed cells were then centrifuged at 20,000 x g in an HB-4 rotor for 10 minutes at 4°C to separate the soluble and insoluble fractions. The pellet, which contained the fusion protein in an insoluble form, was resuspended in 50 ml 0.1 M Tris-HCl, pH 8.0 and collected again by centrifugation. The cell lysate, which contained the soluble proteins, was further clarified by centrifugation at 200,000 x g for 20 minutes in a Beckman L5-65 ultracentrifuge.

2. Purification of recombinant protein under denaturing conditions

The washed inclusion bodies were resuspended in 10 ml 6M guanidine hydrochloride (GuHCl), 0.1 M Tris-HCl, pH 8.0, and stirred at room temperature for 30 minutes. The solution was centrifuged at 40,000 g for 20 minutes. One milliliter of Ni-NTA-agarose was packed in a disposable mini-column (9 mm I.D.), and equilibrated with 6M GuHCl, 0.1 M Tris-HCl (pH 8.0), 10 mM β-mercaptoethanol (β-ME). The supernatant was applied on the NTA-agarose column at a flow rate ~ 0.3 ml/minute. The column was washed sequentially with the following solution: 3 ml of 6M GuHCl, 0.1 M Tris-HCl (pH 8.0), 10 mM β-ME; 3 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 6.5), 10 mM β-ME. Elution was performed by the sequential addition of the following solution: 2 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 mM β-ME; 4 ml of 6M GuHCl, 0.1 M sodium phosphate (pH 5.9), 10 m
phosphate (pH 4.5), 10 mM β-ME; 2 ml of 6M GuHCl, 0.2 M acetic acid. Fractions of 1 ml volume were collected, spotted on a Whatman GF/A glass fiber filter, and stained with Coomassie BR-250 to detect the peak fraction. Alternatively, 6M urea was used in the wash and elution buffer instead of guanidine hydrochloride, which allowed the direct analysis of sample by SDS-PAGE.

3. Purification of the recombinant protein under native conditions

Imidazole was added to the clarified cell lysate to a final concentration of 2 mM, then 0.5 ml of Ni-NTA-agarose, equilibrated with sonication buffer, was added and stirred at 4°C for 1 hour. The NTA-agarose was then transferred to a microcolumn and washed sequentially with the following: 5 ml of sonication buffer with 2 mM imidazole, 10 ml of sonication buffer with 20 mM imidazole, 10 ml of sonication buffer with 50 mM imidazole. Fractions of 0.5 ml were collected for the first 2 ml of each change of the washing solution. The bound fusion protein was eluted with sonication buffer containing 125 mM imidazole, and collected in 250 μl aliquots. The wash and elute fractions were analyzed by SDS-PAGE.

D. STUDIES OF THE PROPERTIES OF THE SOLUBLE RECOMBINANT PROTEIN

1. Sucrose density gradient centrifugation

A 5% - 20% sucrose density gradient (Gibbons and Fronk, 1979, Tang et al., 1982) was prepared in sonication buffer in a 4 ml centrifuge tube for Beckman SW 60 rotor; then 200 μl sample were overlaid on top of the gradient, and centrifuged at 32,000 rpm for 12 - 13 hours at 6.5°C. Catalase (S_{o}^{"}=10.3)
was used as standard protein for the determination of sedimentation coefficient. The gradient was retrieved with a peristaltic pump in 200 μl aliquots and analyzed with SDS-PAGE.

2. **Vanadate mediated photo-cleavage**

Soluble fusion protein was dialyzed against several changes of photo-cleavage medium (50 mM NaAc, 20 mM HEPES, pH 7.5, 1 mM DTT, 0.2 mM EDTA). Protein samples supplemented with different combinations of vanadate, cations and nucleotides were irradiated in a microchamber with an EN-28 lamp (Spectronics Corp, Westbury, NY) as described (Gibbons et al., 1987; Tang and Gibbons, 1987). The microchamber, consisting of the inside of a cap of a 1.5 ml microtube and a cover glass, was maintained at 0°C during the irradiation. The irradiated samples were separated on 13 % SDS-PAGE, and stained with Coomassie Blue, destained, and enhanced by silver nitrate staining (Schoenle et al., 1984).

3. **Effect on the motility of reactivated sea urchin sperm**

The effect of the recombinant proteins on the motility of reactivated sea urchin sperm was examined by the methods previously described (Gibbons, B. H. and Gibbons, 1979, 1976, 1973). Briefly, 25 μl of a sperm suspension from *Colobocentrotus atratus* was added to 0.3 ml of demembranating-extraction solution (0.5 M KCl, 4 mM CaCl₂, 0.1 mM EGTA, 1 mM DTT, 0.04% Triton X-100, 10 mM Tris-acetate buffer, pH 8.1). The extraction was terminated after 50 to 60 seconds by diluting 2 μl of the sperm suspension into a petri dish containing 2.5 ml of reactivating solution [0.2 M potassium
acetate, 2 mM magnesium acetate, 0.1 mM EGTA, 1 mM DTT, 2% PEG 20,000 (polyethylene glycol, w/v), 10 mM Tris-acetate, pH 8.1] and an aliquot of recombinant protein. The mixture was incubated at room temperature for 10 minutes before 5 μl of 21 S dynein (0.5 mg/ml) were added and incubated for 10 minutes; then 1 mM ATP was added. Sperm were observed at 21°C by dark-field light microscopy and the beat frequency of the sperm was measured with a stroboscopic flash unit (Chadsick-Helmuth Co., Monrovis, California).

4. Microtubule co-sedimentation

A procedure similar to that used for the microtubule binding assay of kinesin (Yang et al., 1990) was used. Microtubules (1.8 mg/ml) were assembled from phosphocellulose-purified tubulin (a gift from Dr. Ronald D. Vale) in the presence of 4 μM taxol in PEMG buffer (0.1 M PIPES, pH 6.9, 1 mM EGTA, 1 mM MgCl₂, 2 mM GTP, 1 mM β-mercaptoethanol) at 37°C for 20 minutes. An aliquot of assembled microtubules was then mixed with the protein sample in a total volume of 400 to 500 μl in PEMG, and incubated at room temperature for 30 minutes. The mixture was laid on top of a 100 μl sucrose cushion (15% sucrose in PEMG, supplemented with 0.3 μM taxol) in a microcentrifuge tube (Beckman #344090), and centrifuged at 25,000 rpm (50,000 g) in an SW 65 rotor for 15 minutes at 22°C. After centrifugation, 25 μl of the top supernatant was saved for future analysis, and the rest of the supernatant was removed completely by aspiration. The pellet was dissolved in 1x SDS sample buffer and analyzed by SDS-PAGE.
5. Circular dichroic study

The protein sample was dialyzed extensively against 5 mM phosphate buffer at pH 7.5. The dialysis bag was then sprinkled with dry Sephadex G-50 to absorb water from the protein solution. The concentrated sample was retrieved and centrifuged at 50,000 x \( g \) for 30 minutes to remove any insoluble matter. The protein concentration was determined spectrophotometrically based on its absorbance at 280 nm, and molar extinction coefficient calculated from its amino acid composition by the computer program "PHYSCHEM" in "PC/GENE". The protein sample was placed in a 0.1 mm path-length cell, and circular dichroic (CD) spectra were recorded at 1 nm intervals over the wavelength range 190-300 nm, using a Cary Model 61 spectropolarimeter at room temperature. The CD data were expressed as the mean residue ellipticity, \( \theta \), in units of degree centimeter squared per decimole, using mean residue weight obtained from the amino acid composition. The measured data points covering the range 190-240 nm were analyzed after smoothing with a five-point cubic least-squares filter. Estimates of the apparent secondary structure were computed by using a least-squares curve-fitting procedure to reference spectra derived from published data for standard proteins (Mocz and Gibbons, 1990).
CHAPTER III
RESULTS

A. ISOLATION AND CHARACTERIZATION OF cDNA CLONES

1. Synthesis of cDNA and construction of the cDNA library

Sea urchin embryo messenger RNA was converted to cDNA using a genetically engineered Moloney murine leukemia virus reverse transcriptase which lacks the RNase H activity associated with the natural enzyme (Kotewicz et al., 1988). Double stranded cDNA was synthesized by an improved one-tube strand replacement synthesis method based on previously reported methods (Gubler and Hoffman, 1983; D'Alessio and Gerard, 1988; Okayama and Berg, 1982) in the presence of E. coli DNA polymerase I, E. coli DNA ligase and RNase H. The quality of cDNA synthesized was analyzed by electrophoresis through an alkaline agarose gel and visualized by autoradiography.

In the initial cDNA synthesis experiments, oligo (dT)$_{24}$ was also used as primer in a manner similar to the gene-specific primer. cDNA synthesized using either oligo d(T)$_{24}$ or gene-specific primers yielded long strands of cDNA. However, the cDNAs primed with oligo(dT)$_{24}$ had a large quantity of low molecular weight cDNA species (<2 kb) and a few bands appeared below 1 kb (Figure 3.1 A). The existence of these bands probably reflected the heterogenous nature of mRNA species primed by the oligo(dT)$_{24}$, and the bands could represent the abundant mRNA species. The cDNAs primed with oligonucleotides specific to the dynein $\beta$-heavy chain yield
cDNAs of much larger size, which appeared as continuously distributed lanes from about 200 bp to over 15 kb, with the highest density around 5 kb (Figure 3.1 B). Some faint bands also appeared in the low molecular weight region of around 2 kb. No significant difference was observed between the sizes of the first strand cDNA and the second strand cDNA. In order to reduce the background, cDNAs primed with gene-specific primers were used for the synthesis of DNA probes by PCR and construction of libraries.

Two methods were used to insert the cDNAs into cloning vectors. In the first method, double stranded cDNA was digested with restriction endonuclease PstI, and ligated to plasmid vector pBluescript digested with PstI without size fractionation. In the second method, cDNA was ligated to the vector via EcoRI linker. DNA fragments below 1 kb in length were eliminated prior to ligation by fractionation on a Sephacryl H-500 column (Figure 3.2). The ligated DNA was used to transform E. coli DH5α cells by electroporation. The transformants were plated out on LB-ampicillin plates directly without amplification. In all, three libraries were constructed: two libraries prepared by the first method from cDNA primed with PCNI1 or MMNX1, and a library prepared by the second method from cDNA primed with MPNI1 (Table 1). Some $10^4$ to $10^5$ primary transformants per microgram of mRNA were obtained.

2. Screening of the cDNA library and characterization of cDNA clones

A 2.7 kb DNA fragment which represent nt5175 to nt7956
of the coding sequence of β-heavy chain was amplified from single-stranded cDNA by polymerase chain reaction using primers M1NX2 and MRSQ2, and used for the screening of the two libraries primed with PCNI1 or MMNX1. Some twelve primary clones were picked from these two libraries, and seven of them, named MR1 through MR7, passed further screening. Restriction digestion of plasmids isolated from these clones indicated that plasmids pMR1, pMR2, pMR3 and pMR7 all carried a 3.2 kb fragment. Plasmid pMR3 also carried a small ~700 bp fragment. Plasmids pMR4 and pMR6 carried a 2 kb fragment. Plasmid pMR5 carried a 3.1 kb fragment.

The cDNA inserts in these clones were then amplified by PCR, using primers SK and KS located on the plasmid vector. The PCR products were then purified by agarose gel electrophoresis and sequenced using the same primers. Nucleotide sequences obtained from the 5'-ends and 3'-ends of the inserts in clones pMR1, pMR2 and pMR7 aligned with the 5'-end and 3'-end of a predicted 3193 bp PstI restriction fragment from nt6315 to nt9508 in the coding sequence of β-HC. The end-sequences of pMR3 aligned with the 5'-end of the 3199 bp fragment, and sequences aligned near the PstI site at nt10596. Therefore, pMR3 probably contained the 3193 bp fragment and the 792 bp fragment (from nt9804 to nt10596), which may have been generated by random ligation of restriction fragments. The 3'-end and 5'-end sequences of pMR5 matched with the 5'-end of the 3199 bp fragment, but were about 280 bp short of the 3'-end of the predicted PstI site.
Nucleotide sequences from pMR4 and pMR6 aligned with a 1935 bp fragment which started from nt4380 and ended at nt6315.

Insert isolated from pMR1 was labelled and used as the DNA probe for the screening of the library primed with MPNI3, which yielded 25 cDNA clones. Plasmids were isolated from these clones, and the sizes of the cDNA inserts they carried were determined by restriction digestion with ApaI and NotI, which released the inserts, followed by agarose electrophoresis. The size of the inserts ranged from 1 kb to 13 kb, but most of them were between 5 - 6 kb. The identities of the inserts in several clones were further analyzed by digestion with EcoRI (Figure 3.3).

Nucleotide sequences at both 5’- and 3’- end of the inserts were obtained from plasmids pN12, pN3, pJ3.1 and pJ3.3 by sequencing the PCR-amplified inserts, or ApaI-linearized plasmids. The sequences from one end obtained with the SK primer in the plasmids pN12, pJ3.1 and pN3 all aligned with the 3’- end of the \( \beta \)-HC coding sequence where the MPNI3 primer was placed, while the opposite ends obtained with the KS primer extended to nt7029, nt4579 and nt4147 on the \( \beta \)-HC coding sequence respectively. Therefore, the sizes of the insert they carried should consist of 6581 bp, 9033 bp and 9465 bp of cDNA respectively. These values agreed well with the ones deduced from agarose gel electrophoresis. Nucleotide sequence of the plasmid pJ3.3 aligned with nt441 (from KS primer) and nt11320 (from SK primer) of the \( \beta \)-HC coding sequence. The deduced insert size in pJ3.3 was therefore 10,880 bp. This value was about 2 kb smaller than that
indicated by the gel electrophoresis result (Figure 3.4), probably as a result of deletions during cloning. It is worth noting that the orientation of all four inserts were such that the 3'-end of the \( \beta \)-heavy chain coding sequence was located towards the upstream region of the \( \beta \)-galactosidase gene. Transcription initiated from \( \text{lacZ} \) gene would transcribe the antisense strand of \( \beta \)-heavy chain cDNA.

Plasmids J3.1 and pN3 were sequenced again using the T3 primer which was further upstream of the \( \beta \)-galactosidase gene. It was found that the NotI site immediately following the stop codon in primer MPNI3 was preserved, although about 8 bases that followed the NotI site were lost during cloning.

Overall, about 1% to 5% of the unamplified primary transformants reacted positively with either probe. But on average, only about 25% of the primary picks survived additional rounds of nucleotide probe screening. Overall, about 120 primary clones were picked, and 33 independent clones carrying dynein \( \beta \)-heavy chain cDNA inserts were isolated. The size of the characterized inserts ranged from 1 kb to over 10 kb. Inserts of 5 - 6 kb were common in the EcoRI library, coincided with the distribution pattern of cDNAs electrophoresed on agarose gel (Table 2).

3. Assembly of a near full-length cDNA clone.

The strategy for the assembly of a nearly full-length dynein \( \beta \)-heavy chain cDNA clone is illustrated in Figure 3.5. Plasmid pJ3.1 was digested with ApaI and NsiI, and the resulting 10.58 kb fragment was isolated. Plasmid pJ3.3 was digested with ApaI, NsiI and SmaI, and the resulting 5.33 kb
NsiI/ApaI fragment was isolated. These two DNA fragments were ligated and introduced into \textit{E. coli} DH5\(\alpha\) cells by electroporation. ApaI/NotI digestion of the plasmid isolated from the transformant produced a DNA fragment about 14 kb in length and a 3 kb DNA fragment which is the pBluescript arm. Verification of the completed construct by EcoRI digestion produced DNA fragments of 1.4, 1.8, 2.0, 3.4 and 4.4 kb respectively, along with the 3 kb vector arm. This pattern agreed with that predicted for a cDNA sequence that encodes the dynein \(\beta\)-heavy chain (Figure 3.6).

**B. STABLE EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS: THE LARGE FRAGMENTS**

One of the goals of cloning is to express the cloned cDNA in order to obtain recombinant protein. The most stringent application of the constructed cDNA clone is to use it for protein expression, since any mishaps that disrupt the reading frame of the insert will result in premature termination of translation. Long strands of cDNA that encode major portion of the dynein \(\beta\)-heavy chain were expressed, in part, for this purpose.

1. **Expression of recombinant protein in pET5 vectors**

The pET vectors (Studier and Moffatt, 1986) are one of those systems designed for high level expression of cloned target genes under the control of strong bacteriophage T7 transcription and translation signals, the \(\phi10\) promoter and the translation initiation signals for gene 10 protein (the major capsid protein of T7). For protein production, the established plasmid is transferred into an expression host.
that supplies T7 RNA polymerase. The *E. coli* host BL21(DE3) is a lysogen of bacteriophage DE3, a derivative of λ which contains the T7 RNA polymerase gene under the control of the inducible lacUV5 promoter. Addition of IPTG to a growing culture induces the polymerase, which in turn transcribes the target DNA in the plasmid.

cDNA inserts from plasmids pMR1, pMR4 were isolated by PstI digestion. cDNA inserts in plasmids pN3, pJ3.1 and pN12 were rescued by digestion with ApaI and SmaI. These DNA fragments were treated with T4 DNA polymerase to remove protruding 3′-termini. The expression vector pETS in three different reading frames was digested with EcoRI, and then treated with Klenow enzyme to fill-in the recessive 3′-termini. The DNA insert and the expression vector were ligated, and introduced into the *E. coli* host BL21(DE3). The transformants were screened with rabbit antiserum or monoclonal antibodies against dynein β-chain to isolate clones expressing recombinant dynein protein. Small scale liquid cultures of the antibody-positive clones were induced with IPTG to express the recombinant proteins. SDS-PAGE and Western blot analysis of these mini-cultures indicated the presence of large polypeptides that reacted with anti-dynein antibodies. The size of the DNA inserted into the expression construct correlated well with the apparent molecular weights of those polypeptides, suggesting that the cDNA inserts in these clones do not contain any interruption in their reading frame or major deletions (Figure 3.7 A).

However, problems emerged in the attempt to obtain
homogenous expression clones and to scale up the culture. Antibody staining of the nitrocellulose blot of a replated mini-culture from a well isolated antibody-positive clone would often show a heterogenous population of colonies: with some reacting strongly with the antibody, but the majority only reacting weakly (Figure 3.7 B). Western blot analysis of the time course of induction indicated that these expression clones, in fact, constantly expressed the recombinant proteins even in the absence of the inducer (Figure 3.7 A). This lead to a serious stability problem when a large insert was carried in the expression construct.

2. Expression of the recombinant protein in pET24 vectors

To solve this "leaky expression" problem, pET24 vectors (Novagen, Inc.) were used as the vehicle for the stable expression of dynein recombinant proteins. Unlike the native T7 promoter in pETS, pET24 has a hybrid type of promoter which consists of T7 promoter immediately followed by a lac repressor binding site. The gene that encodes the lac repressor (lacI) was also inserted into the vector. The lac repressor acted both at the lacUV5 promoter to repress transcription of the T7 RNA polymerase by the host RNA polymerase, and at the T7lac promoter to block transcription of the target by a low level of T7 RNA polymerase.

The 4425 bp EcoRI fragment that encoded about one-third of the carboxyl terminal portion of the dynein \( \beta \)-heavy chain (roughly equal to the 130 kDa fragment) was isolated from cDNA clone pN12 and subcloned into the EcoRI site in the pET24a vector. The ligated DNA was introduced into the \( E. \) coli host
BL21(DE3), and expression clone ER1470/BLDE was isolated by antibody screening. The basal expression level of the recombinant protein in this clone was non-detectable in the absence of inducer. Recombinant protein R1470 of the expected molecular weight of 173 kDa was produced upon induction with IPTG. However, the recombinant protein was deposited as insoluble aggregate (the "inclusion bodies") inside the cells. Growth and induction of the expression clone at lower temperature (22°C) did not have any appreciable effect on the solubility of the recombinant protein (Figure 3.8). ER1470/BLDE was co-transformed with pGroESL (Goloubinof, Gatenby and Lorimer, 1989), a compatible plasmid that expresses the E. coli chaperon 60, and grew in a medium that selected for both plasmids. Growth and induction of this expression clone, ER1470/GroESL/BLDE, at 22°C did not improve the solubility of the truncated dynein protein (Figure 3.9). Other conditions that were thought to influence the solubility of recombinant protein upon cell lysis, such as salt concentration, the presence of non-ionic detergent (Triton X-100), polyethylene glycol and substrate (ATP), did not have any effect at all.

Plasmid pER1470 was isolated from clone ER1470/BLDE, digested with BamHI, and the 1532 bp fragment which contained the EcoRI/BamHI fragment (nt8976 to nt10502) of the β-heavy chain sequence and the BamHI fusion cloning site from pET24 was isolated. Ligation of this DNA fragment to BamHI digested pET24 yield plasmid pRB512. The recombinant protein encoded by this construct contained the region between amino acid
residue 2993 to 3503 of the β-heavy chain sequence, as well as the first 11 amino acid residues from T7 major capsule protein at its N-terminal and polyhistidine at its C-terminal. The plasmid was introduced into BL21(DE3), and the clone RB512/BLDE was isolated by screening with anti-β-heavy chain antibodies. Induction of RB512/BLDE culture with IPTG yielded a 60 kDa protein that reacted positively with anti-β-heavy chain antibody and monoclonal antibody 6.31.24 (Piperno, 1984) (Figure 3.10). Inclusion bodies were harvested from induced RB512/BLDE culture, dissolved in 6M GuHCl and applied to Ni-chelating chromatography. Recombinant protein RB512 was purified to 80 - 90% homogeneity and refolded by dialysis against step-wise decreased concentrations of GuHCl, which yielded soluble RB512.

Several other large fragments of dynein β-heavy chain cDNA were also subcloned and expressed with the pET24 vector system. These clones covered essentially all the cloned cDNA sequence in pD1.3 as illustrated in Figure 3.11. Nearly all of the insertions were carried out by ligating the blunt-ended cDNA fragment with a blunt-ended pET24 vector. Induction of BL21(DE3) carrying these plasmids yielded insoluble recombinant proteins of molecular weight comparable to calculated values based on the insert sizes, as visualized by Coomassie staining or by Western blot analysis (Figure 3.12).

3. Refolding of AN1480 and P748

Similar to R1490 and RB512, newly expressed recombinant protein AN1480 (amino acid residue 1208-2688) and P748 (amino acid residue 1358-2107) were deposited in inclusion bodies in
bacterial cells. These two recombinant proteins overlapped with a 92-kDa proteolytic fragment that is relatively resistant to trypsin digestion under mild denaturing conditions, such as mild heat (37°C) or 25% methanol (Mocz, Farias and Gibbons, 1991), and also have above-average hydrophobic amino acid compositions.

The recombinant proteins were purified in 6M guanidine hydrochloride (Figure 3.13), and refolded under a variety of conditions. AN1480 remained soluble until GuHCl concentration was reduced to just below 1M. Further reduction in GuCl concentration induced a rapid transition that led to near complete aggregation of the protein. The redissolution of the aggregates required 1.5 - 2M higher concentration of GuHCl. P748 behaved similarly with a slightly higher GuHCl concentration at the transition points. The sharp transition of solubility and difference in GuCl concentrations at the two transition points suggested that extensive secondary structure might have formed in 1M GuHCl, and these structures were stable under higher GuHCl concentrations.

C. EXPRESSION AND PURIFICATION OF NTP BINDING DOMAINS

It was speculated that the difficulties encountered in obtaining large truncated recombinant dynein protein in a soluble form were in part due to the complexities involved in the folding or refolding of large polypeptides. A smaller recombinant protein could have a better chance to fold correctly, and would expose less hydrophobic amino acid side chains to the aqueous environment. Therefore, soluble recombinant protein may be obtainable by reducing its size.
Several clones representing the nucleotide binding sites of dynein β-heavy chain were constructed in the following experiments, and the names and pedigrees of all the expression clones are diagrammed in Figure 3.14.

1. **Identification of the NTP binding domains**

   It is proposed in this study that the four consensus P-loops in dynein heavy chain sequences probably evolved through gene duplication, a very common mechanism of generating repetitive sequences (Creighton, 1984). Therefore one or more of the nucleotide triphosphate binding motifs or domains should still have the ability to fold, independent from other P-loops, into a structure similar to that in the native dynein and would be able to bind nucleotide triphosphate. One way to determine the minimal amino acid sequence that folds into the nucleotide binding domain is to study the sequence homology of dynein heavy chain isoforms.

   Amino acid sequences of dynein heavy chain isoforms, including both cytoplasmic and axonomal ones and from different organisms, were analyzed using the computer alignment program "Clustal" in PC/GENE (IntelliGenetics), written by the method of Higgins and Sharp (1988, 1989). The alignment of these amino acid sequences in the vicinity of the P-loops is presented in Appendix II, and is schematically represented in Figure 3.15. As indicated by the alignment, the P-loops and the surrounding regions were highly conserved sequences, separated by non-conserved regions. Gaps were often inserted in these non-conserved region to achieve optimum alignment. The boundaries of the NTP binding domains
should lie in these less conserved region. The position and length of these domains were deduced so that the continuity of the conserved sequences were preserved and four sets of nucleotide binding domains were able to fit into the sequence of dynein heavy chain. To satisfy these two requirements, 50 to 100 amino acid residues should be placed to the N-terminal of the P-loop, and 150 to 200 amino acid residues should be placed to the C-terminal of the P-loop. These sequences probably represent the minimum number of amino acid residues required to form an NTP binding domain.

2. Cloning and expression of P1 and P2 NTP binding domains in pET24 vector

PCR primers were designed to flank the DNA sequences that encode the putative NTP binding domains. The N-terminal primers were fitted with an NheI site, and the C-terminal primers were fitted with an EcoRI site. The sequences of the primers are listed in Appendix I. DNA fragments corresponding to the putative NTP binding domain were amplified from pD1.3 by PCR, digested with NheI and EcoRI, ligated into NheI/EcoRI digested pET24, and introduced into BL21(DE3). SDS-PAGE analysis of small scale cultures of individual clones identified those producing recombinant proteins upon IPTG induction. The size and the identity of the inserts carried in these constructs were verified by PCR amplification and partial sequencing of the 3′- and 5′-ends.

Three P1 and P2 expression clones were constructed: 349H/BLDE, 272H/BLDE and 231H/BLDE, which express recombinant proteins representing amino acid residues 1721 - 2069, 1798 -
2069 and 2087 - 2317 in dynein β-chain sequence respectively. Recombinant proteins produced by these clones were insoluble and were deposited in inclusion bodies. The mobility of the recombinant proteins on SDS-PAGE was faster than the protein molecular weight standards. As a result, the apparent molecular weights of all three recombinant proteins were about 20% less than the value calculated from their amino acid sequence. However, analysis of the inserts in these clones by PCR did not show any sign of deletion. Further, the ability of recombinant proteins to binding to Ni-chelating resin through the C-terminal polyhistidine excluded the possibility of premature termination during protein translation.

The insoluble recombinant proteins were dissolved in 6M GuHCl and purified by Ni-chelating chromatography. Refolding was attempted by dialyzing the recombinant protein in 6M GuHCl against a gradually decreasing concentration of GuHCl over a period of 48 to 72 hours in the presence of MgATP (0.5 - 1 mM), reduced/oxidized glutathione (3 mM:0.3 mM), and sodium pyrophosphate (0.5 mM). Most of the recombinant protein precipitated as the denaturant concentration decreased. Only trace amounts of soluble proteins were recovered from the dialysate.

An on-column refolding procedure was used to refold purified recombinant protein. Recombinant protein 349H was absorbed onto the Ni-NTA-agarose matrix in 6M GuHCl at pH 8.0, contaminant proteins were washed away when the pH of the buffer was reduced to pH 6.0, then the pH was increased again to pH 8.0. A linear gradient of GuHCl from 4M to zero was
pumped through the matrix over a period of 94 hours. The soluble protein was eluted with 25 mM EDTA, and the insoluble protein was dissolved and eluted with 6M GuHCl. About 5-10% of the recombinant protein adsorbed on the NTA-agarose eluted as soluble protein. The eluted protein fractions were analyzed by SDS-PAGE (Figure 3.16). About 30% of the recombinant protein eluted from the column appeared intermolecularly cross-linked. This cross-linking was probably caused by the oxidation of sulfhydryl groups but could not be reversed by β-mercaptoethanol or dithiothreitol. Recombinant protein 349H solubilized with this method exhibited no substrate binding capability.

3. Cloning and expression of the NTP binding domains as thioredoxin or glutathione-S-transferase fusion proteins

The plasmid expression vector pTrxFus (Invitrogen) was designed to express a cloned gene as a fusion protein with E. coli thioredoxin to improve its solubility (LaVallie et al., 1993). An 11 kDa protein coded by the trx A gene, thioredoxin folds efficiently and remains soluble even at expression level of up to 40% of the total cellular protein. It also has very high thermal stability with Tm ~ 85°C. These properties may keep folding intermediates of linked heterologous proteins in solution long enough for them to adopt their correct final conformation. Thioredoxin and soluble thioredoxin fusion protein may also be transported to adhesion zone, which can be extracted by osmotic shock.

Plasmid p272T was constructed by excising the NheI/XhoI fragment of p272H, filling-in with Klenow enzyme, and ligating
to pTrxFus which was predigested with BamHI and filled-in with Klenow enzyme. *E. coli* strain GI724 was transformed with p272T to yield clone 272T/GI724. Several clones were cultured in tryptophan-deficient medium, then induced with the addition of tryptophan at 15°C for 16 hours. Cells were harvested and subjected to osmotic shock to extract proteins that were transported to the adhesion zone. Analysis of the fractionated sample by SDS-PAGE revealed large quantities of thioredoxin fusion protein 272T in the pellet fraction, instead of the osmotic shock fluid (Figure 3.17).

Glutathione-S-transferase (GST) is a common 26-kDa cytoplasmic protein of eucaryotes. The GST gene from *Schistosoma japonicum* was used to construct the pGEX vectors that express the target gene that fused to the C terminus of GST (Smith, 1993) in a strategy similar to that of the pTrxFus vector. Plasmid p272GST was constructed by amplifying the sequence that coded for the 272 amino acid residue region in P1 with primers 1798B and 2069E, using pD1.3 as the template. The PCR product was digested with BamHI and EcoRI, ligated into BamHI/EcoRI digested pGEX-4T-1 vector (Pharmacia). *E. coli* strain JM101 was transformed with the expression plasmid to yield clone 272GST/JM101. The clone was cultured and induced with IPTG at 15°C for 3 hour. The cells were harvested, lysed and fractionated. Analysis by SDS-PAGE indicated that 272GST fusion protein was present only in the pellet fraction, but not in the supernatant fraction (Figure 3.18).

The DNA fragment that encodes *E. coli* thioredoxin was
obtained by PCR amplification using primers TRX-NHE and TRXlETK. The PCR product was cloned into pBluescript to yield plasmid pTRX/SK. This plasmid was sequenced with SK and KS primer to verify the authenticity of the thioredoxin gene it carried. Then the thioredoxin gene was excised from the plasmid, and inserted into the NheI site of expression plasmids p349H, p272H and p231H to yield the expression plasmids p349TH, p272TH and p231TH respectively. These expression constructs would express dynein β-HC fragment with thioredoxin at their N-terminal and poly histidine at their C-terminal. E. coli strain BL21 (DE3) was transformed with these plasmids to yield the expression clones 349TH/BLDE, 272TH/BLDE and 231TH/BLDE. Fusion proteins of predicted size were induced with the addition of IPTG into the culture medium.

The expression clones 349TH/BLDE, 272TH/BLDE, 231TH/BLDE and 272GST/JM101 were co-transformed with the plasmid pGroESL. The transformants were grown at 30°C in a medium selective for both plasmids, and were induced by the addition of IPTG at 14°C for 3 hours. Both GroEL and GroES were expressed in addition to the dynein fusion proteins. However, the dynein fusion proteins remained insoluble.

4. Construction and expression of recombinant proteins in a low copy number plasmid vector

To reduce the expression level of fusion proteins and to render the expression construct compatible with plasmid pJM2 that carries the E. coli chaperon 70 gene (McCarty and Walker, 1991), a set of low copy number expression vectors was

The T7/lac promotor in pET24(a-c) was first replaced with the plain T7 promotor from pET5 by replacing the 110 bp BglII/NheI fragment in pET24 with the 103 bp BglII/NheI fragment from pET5. Plasmids from the modified pET24(a-c) were isolated, digested with DraIII, treated with T4 DNA polymerase, and digested again with BglII. The 580 bp DraIII/BglIII fragments which contained the T7 promotor, multiple cloning sites and polyhistidine were isolated. pACYC184 was digested with BamHI and HincII, the 2902 bp fragment which contained the pAl5 replication origin and chloramphenicol resistance gene was isolated, ligated with the 580 bp DraIII/BglI fragments to yield the 3482 bp expression plasmids pHRI(1-3) (Figure 3.19).

DNA fragments that code for dynein or thioredoxin-dynein fusion proteins were excised from p349H, p272H, p231H, p349TH and p231TH by NdeI/XhoI digestion, ligated to NdeI/XhoI digested pHRI to yield plasmids pR349H, pR272H, pR231H, pR349TH and pR231TH respectively. Strain BL21(DE3) was transformed with these low copy number plasmids, followed by co-transformation with pJM2, which contained the gene of E. coli DnaK protein (chaperon 70) and its functional partner, DnaJ, under the control of lac promotor. The transformants were grown in medium containing chloramphenicol and ampicillin which select for both plasmids. The resulting expression clones, R349H/JM2/BLDE, R231H/JM2/BLDE, R349TH/JM2/BLDE and R231TH/JM2/BLDE, constitutively expressed large quantities of
chaperon proteins and small quantity of recombinant dynein proteins. The level of the chaperon protein could be as high as 10 - 20% of the total cellular protein.

The recombinant dynein protein (349H and 231H) remained completely insoluble even in the presence of a large quantity of chaperon 70. Thioredoxin-dynein fusion proteins (349TH and 231TH) were present in the soluble fraction of cell lysate in small quantities. The fusion protein in the lysate was absorbed on Ni-NTA-agarose, washed with an imidazole step gradient at pH 8.0 and eluted with 125 mM imidazole (Figure 3.20). Most of the E. coli proteins and chaperons were removed by washing with 50 mM imidazole. A small quantity of chaperon 70 and an unidentified protein co-purified with the fusion protein. This unknown protein of 33 kDa associated with the fusion proteins in a nearly 1:1 ratio, and therefore, was probably a host protein. Typically, between 2 to 3 mg of fusion protein were obtained from 500 ml of liquid culture after two rounds of binding on 0.5 ml Ni-NTA-agarose.

D. PROPERTIES OF SOLUBLE RECOMBINANT PROTEINS

1. The aggregation status of the soluble fusion protein

The aggregation state of the fusion protein 231TH was analyzed by ultracentrifugation over a 5 - 20% sucrose gradient. The fusion protein band had a relatively sharp rear edge, but the leading edge extended from about 10S towards the bottom of the gradient, which suggested that the fusion protein was partially aggregated. The unknown 33 kDa E. coli protein was associated with the fusion protein throughout the gradient (Figure 3.21).
2. Vanadate mediated photo-cleavage

Fusion protein 349TH irradiated under the standard V1 (Gibbons et al. 1987) or V2 (Tang and Gibbons, 1987) cleavage conditions for dynein heavy chains did not show any sign of cleavage. Photo-cleavage was observed with millimolar concentrations of vanadate, ADP and prolonged irradiation (2-10 hours) (Figure 3.22). Several bands absent in the pre-irradiated sample appeared after irradiation, and intensified with the increase of irradiation time and vanadate concentration. Therefore, the cleavage condition observed here displayed certain characters of the V2 cleavage of native sea urchin axonemal dynein, which is favored by high vanadate concentration, divalent cations (Mg$^{2+}$ or Mn$^{2+}$), and absence of ATP. Although it was unable to positively link the degradation of 349TH and the appearance of the lower bands, due to the unsatisfactory purity of the fusion protein and the lack of proper antibodies available to stain the fusion protein, the apparent molecular weights of three of these bands, 27-, 25- and 22-kDa, were very close to the expected cleavage products, if the photo-cleavage happened in the vicinity of the P-loop of the fusion protein. Such a cleavage, similar to that in native dynein (Gibbons et al., 1991), should generate two peptide fragments of Mr 28-kDa and 26-kDa. Further experiments with different combinations of nucleotides (ATP or ADP) and divalent cations (Mg$^{2+}$ or Mn$^{2+}$) indicate that the cleavage condition of the fusion protein was distinct from that occurring in native dynein heavy chains (Figure 3.23).
3. **Effect on the motility of reactivated sea urchin sperm**

The beat frequency of sperm flagella is affected by the number of dynein arms interacting with microtubules in the axoneme. If the fusion protein competes with native dynein for binding to microtubules, a reduction of the beat frequency could be expected. Alternatively, the competition could interfere with the restoration of beat frequency by 21S dynein in KCl extracted sea urchin sperm.

Incubation with the fusion protein 349TH preparation did not significantly affect the beat frequency of demembranated sea urchin *Colobocentrotus atratus* sperm, or the beat frequency of sperm with the outer arm dynein removed by KCl extraction. In KCl-extracted sperm, pre-incubation with the fusion protein preparation slightly inhibited the restoration of beat frequency by *Tripneustes gratilla* 21S dynein. The inhibitive effect also increased slightly as the quantity of fusion protein increased (Figure 3.24). In addition, the inhibitive effect was gradually reversed over a period of 10-15 minutes. Pre-incubation of KCl-extracted sperm with the refolded RB512 (36-72 ug/ml) did not block the restoration of beat frequency by 21S dynein (1 ug/ml).

4. **Effect on in vitro polymerized microtubules**

To investigate whether fusion protein directly interacts with microtubules, recombinant protein preparation was incubated with taxol stabilized microtubules, followed by sedimentation through a 15% sucrose cushion containing GTP and taxol. The refolded RB512 (0.9mg/ml final concentration) did not exhibit any meaningful association with microtubules under
the conditions that enable the binding of native dynein to microtubules, nor did it affect the stability of microtubules. Fusion protein 349TH preparation aggregates extensively under this assay condition. Most surprisingly, polymerized microtubules became disintegrated in the presence of the either 349TH (0.4mg/ml) or 231TH (0.05mg/ml) preparations. This effect was probably not caused by the fusion proteins themself, rather, it could stem from impurities in the preparations. This effect apparently did not depended on the concentration of fusion protein (231TH or 349TH).

5. Circular Dichroism spectra

A circular dichroism (CD) spectrum of refolded RB512 was measured (Figure 3.25). The secondary structure of the refolded protein was computed by using a least-squares curve-fitting procedure comparing with reference spectra derived from published data for standard proteins (Mocz and Gibbons, 1990). The computed data indicated that refolded RB512 contained about 25% α-helix, 28% β-form, and a total of about 82% secondary structure components. These values were similar to that of a typical globular protein (Yang et al., 1986).
Table 1. Summary of cDNA libraries

<table>
<thead>
<tr>
<th>Primer for cDNA synthesis</th>
<th>Method of cloning</th>
<th>Probe used for screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCNI1</td>
<td>PstI digestion of cDNA</td>
<td>RT-PCR probe</td>
</tr>
<tr>
<td>MMNX1</td>
<td>PstI digestion of cDNA</td>
<td>RT-PCR probe</td>
</tr>
<tr>
<td>MPNI3</td>
<td>EcoRI linker addition</td>
<td>pMR1 insert</td>
</tr>
<tr>
<td>Clone</td>
<td>Insert size (kb)</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>MR1</td>
<td>3.193</td>
<td></td>
</tr>
<tr>
<td>MR2</td>
<td>3.193</td>
<td></td>
</tr>
<tr>
<td>MR3</td>
<td>3.985</td>
<td></td>
</tr>
<tr>
<td>MR4</td>
<td>1.935</td>
<td></td>
</tr>
<tr>
<td>MR5</td>
<td>2.919</td>
<td></td>
</tr>
<tr>
<td>MR6</td>
<td>1.935</td>
<td></td>
</tr>
<tr>
<td>MR7</td>
<td>3.193</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>9.465</td>
<td></td>
</tr>
<tr>
<td>N6</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>N12</td>
<td>6.583</td>
<td></td>
</tr>
<tr>
<td>N14</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>N15</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>N16</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>N19</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>N21</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>N23</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>J1.5</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>J1.6</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>J1.7</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>J2.6</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>J2.9</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>J3.1</td>
<td>9.033</td>
<td></td>
</tr>
<tr>
<td>J3.3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>J3.4</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>J4.3</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>J5.9</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>J6.1</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>J6.9</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>J6.10</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>J7.1</td>
<td>7.0</td>
<td></td>
</tr>
</tbody>
</table>

Table II. Summary of cDNA clones
Figure 3.1  Autoradiograph of cDNAs separated on agarose gel.
(A) Oligo d(T)$_{24}$ primed cDNA separated on 1% alkaline agarose gel: second strand synthesis.
(B) MPNI3 primed cDNA separated on 1% alkaline agarose gel: lane 1, first strand synthesis; lane 2, second strand synthesis.
Figure 3.2 Chromatography of cDNA on Sephacryl H-500 column

Linkered cDNA was fractionated on a 2-ml Sephacryl H-500 column. 50 μl aliquots were collected and counted in tritium channel in Beckman LS 7000. The cumulative elution volume was plotted against cpm. DNA fragments bigger than 1 kb in length were eluted in the void volume.
Figure 3.3  *EcoRI* restriction digestion analysis of plasmids.

Plasmid DNA was isolated from 1 ml mini-culture of plasmid cDNA clones. Between 1 - 2 ug of DNA was typically digested with 20 units of *EcoRI* in a 15 ul reaction at 37°C overnight. Portions of the digested DNA samples were separated on 1% LE agarose with 0.25 ug/ml ethidium bromide in 0.5 x TBE buffer at 5 V/cm.

Lane 1, pN1;
lane 2, pN3;
lane 3, pN6;
lane 4, pN12;
lane 5, pN15;
lane 6, pJ3.1;
lane 7, pJ3.3;
lane 8, pJ3.4;
M, molecular weight standards (λ DNA-*Bst*E II digest).
Figure 3.4  ApaI/NotI restriction digestion analysis of plasmids. Plasmid DNA from mini-preparation was digested with 20 units of ApaI in a 10 ul reaction in NEBuffer 4 for 2 hours at 37°C, followed by digestion with 20 units of NotI at 37°C overnight. The digested DNA samples were separated on 1% LE agarose with 0.25 ug/ml ethidium bromide in 0.5 x TBE buffer at 5 V/cm. 

Lane 1, pN1; 
lane 2, pN6;  
lane 3, pJ3.1; 
lane 4, pJ3.3;  
M, molecular weight standards (λ DNA-BstE II digest).
Figure 3.5  Schematic diagram for the assembly of pD1.3.

Plasmid pJ3.1 was digested with ApaI and NsiI, the 3.5 kb fragment was replaced with the 6.8 kb ApaI/NsiI fragment from plasmid pJ3.3. The orientation of dynein insert is in such a way that the stop codon is closest to the SacI site in the multiple cloning site (MCS) on the vector. The numbers in the parenthesis indicates the corresponding nucleotide numbers in the heavy chain coding sequence.
(Dynel nt 233) EcoRI ApaI

SmaI EcoRI NsiI J 3,3 EcoRI KpnI

SmaI EcoRI (Dynel nt 13401) NsiI J 3,1 KpnI

SacI SacI

pD 1.3

M C S lac F O1

CIE E ColI Am
Figure 3.6 Restriction digestion analysis of pDl.3

Between 1 - 1.5 ug plasmid DNA was digested with 10 units of restriction enzyme. One-fifth of the DNA sample was separated on 1% LE agarose gel with 0.25 ug/ml in 0.5 x TBE buffer at 5 V/cm. Lane 1, ApaI/NotI digestion of pDl.3; lane 2, EcoRI digestion of pDl.3; M, molecular weight standards (λ DNA-BstE II digest).
Figure 3.7  Analysis of expression clone e12.

(A) Western blot analysis of expression of recombinant protein in e12.4. Cells harvested from 2 ml of e12.4 overnight culture were used to inoculate 20 ml YT medium with 200 ug/ml ampicillin. The culture was grown at 37°C to O.D. = 0.56, and IPTG was added to 0.5 mM to induce protein expression. The culture was sampled at 0, 15, 45, 75, 105, 130 and 160 minutes following IPTG induction. Samples were separated on 6.5% SDS-PAGE, transferred to nitrocellulose filter, and stained with anti-β-heavy chain antiserum. The blot was developed with AP-conjugated secondary antibody. BLDE, E. coli host BL21(DE3). M, tryptic digested dynein.

(B) Colony-blot of e12.9.1 and e12.10.1 streaked on LB-ampicilline plate stained with anti-β-heavy chain antiserum. Showing heterogeneous staining of bacterial colonies.
Figure 3.8 Analysis of ER1470/BL21(DE3) low temperature induction.

ER1470/BL21(DE3) culture was grown to O.D. = 0.5, then induced with 0.5 mM IPTG at 22°C and sampled at one hour interval for 7 hours. The samples were electrophoresed on 6.5% SDS-PAGE, transferred to nitrocellulose filter and stained with anti-β-heavy chain antiserum. (A) Ponceau S stain of nitrocellulose blot, supernatant;

(B) Ponceau S stain of nitrocellulose blot, pellet; the position of induced R1470 band was marked by an arrow.

(C) Anti-β-HC antiserum stain of blot, supernatant;

(D) Anti-β-HC antiserum stain of blot, pellet;

M, molecular weight standards (dynein trypsin digest).
Figure 3.9 Analysis of pER1470/pGroESL/BL21(DE3) induction

ER1479/GroESL/BL21(DE3) was induced at 22°C for 16 hours, then lysed with (lanes 2, 4, 5, 6) or without (lanes 1 and 3) 13 mM ATP. The lysis buffer also contains 1.2% Triton X-100, 30% sucrose (lanes 1, 2, 5) or 250 mM NaAc (lanes 3, 4, 6). Samples were electrophoresed on 6.5% SDS-PAGE, transferred to nitrocellulose filter, then stained with anti-β-antiserum.

(A) Supernatant fraction;
(B) Pellet fraction.

M, molecular weight standards (dynein trypsin digest).
Figure 3.10 Western blot analysis of RB512

R1470/BLDE and RB512/BLDE were grown to ~ 1.0 O.D., and induced with 1 mM IPTG for 1.5 hour at 37°C. About 16 ul of induced sample were separated on 8% SDS-PAGE, transferred to nitrocellulose filter, and stained with monoclonal antibody 6-31-24.

(A) Ponceaus S stain of nitrocellulose blot.

(B) Monoclonal antibody stain of blot.

Lane 1, BL21 (DE3);
lane 2, R1470/BLDE;
lane 3, RB512/BLDE.

M, molecular weight standards (dynein tryptic digest).
Figure 3.11 Schematic diagram shows the expression of large sections of the β-heavy chain. The following table lists the large expression plasmids constructed in pET24 and the regions of β-heavy chain represented by them. The recombinant proteins encoded by these plasmids contained 8 - 11 amino acid residues at their N-terminal derived from the T7 major capsule protein, and polyhistidine at their C-terminal.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>represented section in ( \beta)-HC (amino acid residue number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pR1470</td>
<td>2992-4466</td>
</tr>
<tr>
<td>pRB512</td>
<td>2992-3504</td>
</tr>
<tr>
<td>pRB1393</td>
<td>2992-4385</td>
</tr>
<tr>
<td>pXA1144</td>
<td>668 -1208</td>
</tr>
<tr>
<td>pAN1480</td>
<td>1208-2688</td>
</tr>
<tr>
<td>pAX2514</td>
<td>1208-3722</td>
</tr>
<tr>
<td>pNX2295</td>
<td>1427-3722</td>
</tr>
<tr>
<td>pSX2050</td>
<td>1672-3722</td>
</tr>
<tr>
<td>pNN2919</td>
<td>1427-4341</td>
</tr>
<tr>
<td>pP748</td>
<td>1358-2107</td>
</tr>
</tbody>
</table>
Figure 3.12  Induction of large dynein expression clones

LB medium with 50 μg/ml kanamycin in 24-well tissue culture plate was inoculated with a single colony, and shaked at 37°C till O.D. reached ~ 0.6. IPTG was then added to 0.5 mM to induce expression for 1.5 hours. The induced sample was electrophoresed on 6.5% SDS-PAGE, transferred to nitrocellulose filter and stained with anti-β-HC antibodies. The upper most band in each lane is presumably the full-length recombinant protein, and the bands below are probably premature or degraded products. P748 was not stained probably because of the low antigenicity of the highly conserved nucleotide binding domains. The faint staining of RB1393 and XA1144 is caused by low yield of recombinant proteins in these particular isolates, and clones obtained later have much high yield of recombinant proteins.

Lanes 1, 2, RB1393/BLDE;
lanes 3, 4, XA1144/BLDE;
lanes 5, 6, AX2514/BLDE;
lanes 7, 8, NX2295/BLDE;
lanes 9, 10, SX2050/BLDE;
lanes 11, 12, NN2919/BLDE;
lanes 13, 14, P748/BLDE;
M, dynein trypsin digestion.
Figure 3.13  A typical purification of AN1480

AN1480/BLDE was grown to 0.6 O.D. in LB medium with 80 ug/ml kanamycin, and induced with the addition of 1 mM of IPTG at 37°C for 3.5 hours. Inclusion bodies harvested from 250 ml of induced culture were dissolved in 6 M Guanidium hydrochloride, 10 mM β-mcaptoethanol. AN1480 was absorbed on 4 ml Ni-NTA-agarose, followed by washing the column sequentially with the following: 25 ml of 8 M urea, 0.1 M Tris-HCl, pH 8.0, 10 mM BME; 40 ml of 8 M urea, 0.1 M phosphate, pH 6.5, 10 mM BME. The bound protein was eluted with 21 ml of 8 M urea, 0.1 M phosphate, pH 4.5, 10 mM BME; followed by 9 ml of 8 M urea, 0.1 M Tris-HCl, pH 8.0, 20 mM EDTA. The eluate was collected in 3 ml aliquot. 20 ul of each aliquot were mixed with 5 ul of 5 x SDS sample buffer and electrophoresed on 6.5% SDS-PAGE.

Lane 1, inclusion bodies;
lane 2 - 9, elution with 8 M urea, pH4.5;
lane 10 - 13, elution with 8 M urea, 20 mM EDTA;
M, dynein trypsin digestion.
Figure 3.14 The pedigree of the P-loop expression clones

Each expression clone is represented by a rectangular box which contained the plasmids and the host. The two soluble fusion proteins, 349TH and 231TH, are represented by the ellipsoidal circles.
Figure 3.15 Schematic presentation of the alignment of amino acid sequences surrounding the nucleotide-binding sites of dynein heavy chain.

The amino acid sequences surrounding the nucleotide-binding sites were aligned the PC/GENE program "Clustal" using the method developed by Higgins and Sharp (1988; 1989). Six, eight and twelve axonemal and cytoplasmic dynein heavy chain sequences were used to produce the alignments for P4, P1 and P2/ P3 respectively, which was detailed in Appendix II. The alignment is schematically presented below, and the four P-loops are aligned with each other. Continuously conserved sections in the alignment are represented by filled boxes. The non-conserved regions, usually with gaps inserted, are represented by single lines. The sketch is drawn in proportion to the sea urchin dynein β-heavy chain sequence.
Inclusion bodies from induced 349TH/BLDE culture were harvested and dissolved in 6M guanidium hydrochloride, 0.1M Tris-HCl, pH 8.0. 349H was absorbed on 4 mL Ni-NTA-agarose. The resin was then packed in a 1 cm (I.D.) column and washed with 20 ml of 6M guanidium hydrochloride, 0.1M phosphate, pH 6.0, followed by 20 ml of 4M guanidium hydrochloride, 0.1M Tris-acetate, pH 8.3. A linear 4 - 0 M guanidium hydrochloride gradient which consisted of 25 ml of 4M guanidium hydrochloride, 50mM Tris-acetate, pH 8.3, 2mM ATP, 2.5mM MgCl$_2$, 0.5mM Na$_3$P$_2$O$_7$, 2mM reduced glutathione, and 25ml of 40% glycerol, 50mM Tris-acetate, pH 8.3, 2mM ATP, 2.5mM MgCl$_2$, 0.5mM Na$_3$P$_2$O$_7$, 2mM reduced glutathione was pumped through the column at 1 ml/hr rate. After refolding, the column was eluted with 10 ml of 50mM Tris-acetate, 25mM EDTA, followed by 6M guanidium hydrochloride, and collected as 1 ml aliquots. Guanidium hydrochloride eluted sample was diluted 1:100 in water and boiled in a water bath for 5 minutes to precipitate protein before mixing with SDS sample buffer. The samples were electrophoresed on 13% SDS-PAGE.

Lane 1 - 6, eluted with 25mM EDTA (fractions 4 - 9); lane 7, eluted with 6M GuCl (fraction 12); M, molecular weight standards.
Figure 3.17 SDS-PAGE analysis of induced 272T culture
1 ml RM medium (1x M9 salts, 2% casamino acids, 1% glycerol, 1mM MgCl₂) with 100μg/ml ampicilin in 24-well tissue culture plate was inoculated with 50μl overnight culture of 272T/GI724, and grew to ~0.5 O.D. at 30°C. The culture was chilled to 15°C and induced by the addition of 0.1mg/ml tryptophan for 16 hours. The cells were harvested, and soluble proteins that had been transported to periplasm (e.g. thioredoxin and thioredoxin fusion protein) were extracted using cold osmotic shock method: the cells was sequentially extracted once with 150μl of hypertonic solution (20% sucrose, 20 mM Tris-HCl, pH 8.0, 2.5 mM EDTA), twice with 150μl each hypotonic solution (20 mM Tris-HCl, pH 8.0, 2.5 mM EDTA) at 0°C. The osmotic shock fluids (hypotonic fractions), along with the pellet fraction were analyzed on 13% SDS-PAGE.
Lanes 1, 2, pelle fraction of 272T;
lane 3, pellet fraction of ThioFusion vector;
lane 4, 5, 7, 8, osmotic shock fluids of 272T;
Lane 6, 9, osmotic shock fluids of ThioFusion vector;
Mr, molecular weight standard.
Figure 3.18  Analysis of induced 272GST culture

1 ml LB-ampicilin medium in 24-well tissue culture plate was inoculated with a single colony of 272GST/JM101 and grown to ~0.6 O.D. at 37°C. The culture was then chilled to 15°C and induced by the addition of 0.2mM IPTG for 3 hours. The cells were harvested and lysed by a combined action of lysozyme and osmotic shock. The lysate was centrifuged at 16,000 x g for 15 minutes to separate the soluble and insoluble fraction, then each fraction was analyzed on 13% SDS-PAGE.

(A) pellet fraction;
(B) supernatant fraction;
Lane 1: GST vector;
lanes 2-7: 272GST clones
Mr: molecular marker.
Figure 3.19  pHR map

The pHR series of low copy-number expression vectors were constructed from pACYC184, pET24 and pET5 as described in the text. The map shows the location of the following elements in these vectors: chloramphenicol resistance gene (Cm, 219-3042); the origin of replication (Ori, 845-847); and the T7 cloning/expression region (1869-2154). Nucleotide number 1 of pHR-a is the first G of the EcoRI site in the Cm gene.
Cells from 1 ml overnight culture of 349TH/JM2/BLDE was washed twice with 1.4 ml cold distilled water, and inoculated to 500 ml LB medium with 200 ug/ml ampicillin and 50 ug/ml chloramphenicol. The culture was shaked at 30°C for 16 hours. 1.7 g of cell pellet was harvested from the culture, and resuspended in 10 ml of sonication buffer (300 mM sodium acetate, 50 mM Na-phosphate, pH 8.0, 10 mM β-mercaptoethanol) supplemented with PMSF and TAME at 200 ug/ml, lysozyme at 1 mg/ml. The suspension was incubated at 0°C for 30 minutes, sonicated intermittently for 1 minute at 20 watts with a microprobe, and stirred at 0°C for 1 hour in the presence of 10 mM ATP. The lysate was cleared by centrifugation at 200,000 g for 30 minutes. 0.5 ml Ni-NTA-agarose was suspended and incubated with the clarified lysate at 4°C for 1 hour, then transferred to a 0.5 cm I.D. mini-column. The resin was washed sequentially, each time with 10 ml sonicate buffer containing 2, 25, 50 mM imidazol. Soluble 349TH was eluted with sonicate buffer containing 125 mM imidazol. Fractions of 0.3 ml were collected, and 20 ul aliquots were analyzed by 13% SDS-PAGE.

(A) Lane 1, GroESL/BL21(DE3);
    lane 2, 349TH/JM2/BLDE lysate;
    lane 3, column flow-through;
    lanes 4 - 7, 20 mM imidazol wash;
    lanes 8 - 11, 50 mM imidazol wash.

(B) Lanes 1 - 13, 125 mM imidazol elution
    Mr, molecular weight standards.
Figure 3.21  Sucrose gradient centrifugation analysis of 231TH

A 5 - 20 % sucrose gradient in sonication buffer was prepared in a 2.1 ml centrifuge tube for Beckman SW-60. 150 ul of 231TH preparation (~0.2 mg/ml) was loaded on top of the gradient, and centrifuged at 32 K for 12.5 hours at 6.7°C. Catalase (10.5 S) was used as sedimentation standard, and loaded on a parellaly prepared gradient. The gradient was retrived from the bottom in 0.2 ml fractions by the use of a peristatic pump after centrifugation. Samples were electrophoresed on 13% SDS-PAGE, and stained with Coomassie Brilliant Blue R-250 and silver nitrate. The position of catalase in the gradient is indicated by the arrows.
Figure 3.22 Vanadate-mediated photo-cleavage of 349TH

1.5ml 349TH preparation (~0.4mg/ml) was dialyzed twice against 20ml of cleavage buffer (50mM Na-acetate, 20mM HEPES, pH7.5, 1mM DTT, 0.2mM EDTA) supplemented with 1mM ATP, 2mM MgCl₂; followed by 20ml of cleavage buffer supplemented with 2mM ADP, 4mM MgCl₂, and 2mM NaVO₃. 21S dynein was added to 38μg/ml as an internal standard of cleavage. The sample was irradiated in a microchamber at 4°C, and aliquots were taken at different time point for analysis by SDS-PAGE. The samples were separated on 13% gel, stained with Coomassie Blue and silver nitrate. Two principle polypeptides of Mr ~28-Kda and ~26-Kda were expected if the cleavage of 349TH is similar to that of native dynein. Lanes 1, 3, 5, 7 were irradiated with 2mM vanadate for 0, 2, 5, 15 hours respectively; lanes 2, 4, 6, 8 were irradiated with 5 mM vanadate for 0, 2, 5, 15 hours respectively. Mr, molecular weight standards. DHCs, dynein heavy chains; CHCs, cleaved heavy chains.

(A) Hypothetical cleavage site of fusion protein 349TH.

(B) Analysis of cleaved samples.
Figure 3.23 Vanadate mediated photo-cleavage of 349TH under various conditions

1.6 ml of 349TH preparation was dialized against 20 ml of cleavage buffer supplemented with 1 mM ATP, 2 mM MgCl₂ at 4°C overnight, followed by three changes against 40 ml each of cleavage buffer. Sea urchin 21 S dynein was added to 37 μg/ml as an internal standard of cleavage. Vanadate and nucleotide were added as specified on each lane. The samples were irradiated for 10 hours at 4°C, electrophoresed on 13% SDS-PAGE, stained with Coomassie Brilliant Blue and silver nitrate.

lane 1, unirradiated sample;
lane 2, irradiated without vanadate;
lane 3, with 4 mM vanadate;
lane 4, with 4 mM vanadate, 1 mM MgCl₂;
lane 5, with 4 mM vanadate, 1 mM MnCl₂;
lane 6, with 4 mM vanadate, 1 mM MgCl₂, 1 mM ATP;
lane 7, with 4 mM vanadate, 1 mM MgCl₂, 1 mM ADP;
lane 8, with 4 mM vanadate, 1 mM MnCl₂, 1 mM ATP;
lane 9, with 4 mM vanadate, 1 mM MnCl₂, 1 mM ADP;

Mr, molecular weight standards;
DHCs, dynein heavy chains; CHCs, cleaved heavy chains.
Figure 3.24 Effect of 349TH preparation on the motility of reactivated sea urchin sperm

Demembranated *Colobocentrotus atratus* sperms were diluted in 2.5ml reactivating solution, and incubated with various amount of 349TH fusion protein (0.6mg/ml) for 10 minutes, followed by the addition of 21S dynein. The sperms were incubated for another 5 minutes before activation. The beat frequencies of 10-20 sperms were counted and expressed as mean ± standard deviation.

<table>
<thead>
<tr>
<th>No.</th>
<th>Treatment</th>
<th>Beat frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Treated sea water</td>
<td>25.0 ± 1.2</td>
</tr>
<tr>
<td>2.</td>
<td>reactivated</td>
<td>29.0 ± 1.4</td>
</tr>
<tr>
<td>3.</td>
<td>reactivated + 50 ul 349TH</td>
<td>28.0 ± 0.8</td>
</tr>
<tr>
<td>4.</td>
<td>KCl extracted</td>
<td>14.9 ± 0.5</td>
</tr>
<tr>
<td>5.</td>
<td>KCl extracted + 50 ul 349TH</td>
<td>15.0 ± 1.2</td>
</tr>
<tr>
<td>6.</td>
<td>KCl extracted + 5 ul 21S</td>
<td>23.6 ± 1.0</td>
</tr>
<tr>
<td>7.</td>
<td>KCl extracted + 25 ul 349TH + 5 ul 21S</td>
<td>22.5 ± 0.9</td>
</tr>
<tr>
<td>8.</td>
<td>KCl extracted + 50 ul 349TH + 5 ul 21S</td>
<td>20.9 ± 1.3</td>
</tr>
<tr>
<td>9.</td>
<td>KCl extracted + 75 ul 349TH + 5 ul 21S</td>
<td>20.6 ± 0.9</td>
</tr>
<tr>
<td>10.</td>
<td>KCl extracted + 100 ul 349TH + 5 ul 21S</td>
<td>20.5 ± 0.8</td>
</tr>
<tr>
<td>11.</td>
<td>KCl extracted + 150 ul 349TH + 5 ul 21S</td>
<td>19.4 ± 1.2</td>
</tr>
</tbody>
</table>
RB512 was purified from solubilized inclusion bodies by Ni-chelating chromatography in 6 M guanidium hydrochloride. 20 ml of RB512 at 0.75 mg/ml were dialyzed sequentially against 4, 1, 0 molar concentration of guanidium hydrochloride which was prepared in 250 ml of sodium acetate buffer (0.45 M NaAc, 50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 50 mM β-mercaptoethanol). The solutions were equilibrated for 6 - 16 hours between each change. The protein was finally dialized against several changes of 5 mM pH 7.5 phosphate buffer, and concentrated by sprinkling the dialysis bag with dry Sephadex G-50. The concentrated RB512 was clarified by centrifugation at 50,000 g for 30 minutes. The refolded RB512 was placed in a 0.1 mm path-length cell, and CD spectra were recorded at 1-nm intervals over the wavelength range 190-300 nm, using a Cary Model 61 spectropolarimeter at room temperature. The dashed line indicated calculated spectrum that is derived from reference proteins. The solid line indicated measured spectra of RB512.
A. ISOLATION AND CHARACTERIZATION OF cDNA CLONES

The construction of a cDNA clone that encodes a high molecular weight polypeptide is often a tedious procedure which involves isolating several individual partial cDNA clones representing the coding sequence and multiple rounds of ligation and subcloning, such as that used in the cloning of the 565 kDa skeletal muscle ryanodine receptor (Takeshima et al., 1989).

In this study, long strands of cDNAs were obtained as a result of using an RNase H+ reverse transcriptase in an improved strand replacement synthesis reaction. Although no single clone that contains the entire coding sequence of dynein β-heavy chain was isolated, the size of the insert in several of the clones was over 7 kb. The isolation of these large cDNA clones facilitated the assembly of a nearly full-length cDNA clone that encode the dynein β-heavy chain.

Several other factors also played key roles in this success. First, the use of electroporation for transferring nucleic acids into E. coli cells (Calvin and Hanawalt, 1988) facilitated the cloning of large inserts. Since the size of the DNA did not seem to affect the transformation efficiency by electroporation (Ausubel et al., 1994), this technique was particularly valuable for the construction of the dynein cDNA library and large dynein cDNA clones. Secondly, the combined use of gene-specific primers in the cDNA synthesis reaction,
and size fractionation of linked cDNA greatly reduced background in the cDNA libraries. Thirdly, the use of large DNA probes located in the center of the dynein β-heavy chain coding sequence boosted the hybridization signal for slow-growing clones carrying large inserts, and eliminated clones carrying short 3′-end inserts.

Although the isolated cDNA clones and the assembled full-length clone were not sequenced entirely, several lines of evidence support their identity and integrity: (1) restriction digestion patterns of the plasmids isolated from these clones matched with those predicted from the β-heavy chain coding sequence (Gibbons et al 1991); (2) DNA sequences obtained from the 3′- and 5′-ends aligned with the β-heavy chain sequence with over 95% identity; (3) expression of over 10 different recombinant proteins with molecular weight proportional to the size of the cDNA insert; (4) the observed immunological properties of the recombinant proteins were identical to those of authentic dynein β-heavy chain.

Most of the sequence discrepancies occurred at the ends of sequencing ladders, where readings were occasionally difficult. Out of the remaining discrepancies, about 70% were observed at positions where degeneracy existed in the β-chain sequence, and about 30% were silent changes.

B. EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS

The choice of expression host

E. coli was chosen as the host for the expression of the cloned dynein gene. Compared to other eucaryotic expression systems that have been used for stable expression of cloned
genes, such as yeast, insect cell culture or mammalian cell culture systems, E. coli as the host cell has the advantages of having well understood biological and genetic characteristics, easy manipulation and quick growth in inexpensive media. Therefore, it has been widely used to generate proteins of interest and commercial value (reviewed in Marston, 1986). Several proteins involved in cell motility have also been successfully expressed in E. coli, including myosin and its accessory proteins (McNally, et al., 1988; Leinwand, et al., 1989), and kinesin (Yang, et al. 1990; Gilbert and Johnson, 1993). There are drawbacks, however, in using E. coli as the expression host. In many cases, although high levels of expression of the desired polypeptides have been obtained, much of the resultant materials do not fold to their native soluble conformations, rather they tend to form microscopic insoluble aggregates, or inclusion bodies. In addition, proteins expressed in E. coli lack certain post-transcriptional modifications, such as glycosylation, N-terminal acetylation and C-terminal amidation, which are common to many eucaryotic proteins.

There are specific disadvantages in using E. coli as the host for the expression of dynein. First, the large size of β-heavy chain almost certainly requires an array of cellular mechanisms to help its folding. These may not be present in the E. coli cell, or even if they are, they may not interact with dynein heavy chain properly. Secondly, there are no dynein accessory proteins, such as intermediate chains or light chains in E. coli, which may associate with and
stabilize the heavy chain. Thirdly, an almost complete lack of knowledge of the domain structure of dynein makes it difficult to identify prospective expression constructs. However, if only truncated forms are to be expressed, the requirement for cellular folding mechanisms or accessory proteins may be reduced. Furthermore, if the truncation causes difficulties in protein folding or in the stability of the protein, expression in other systems may not be able to alleviate the problem either.

The stability of expression clones

Expression clones constructed in pET5 vectors and BL21(DE3) were extremely unstable. In these clones, low level expression of dynein recombinant proteins persisted even in the absence of IPTG. Lactose, the natural inducer of lac operon which exists in various amounts in culture media containing casein amino acids, may induce low level expression of T7 RNA polymerase under the control of lacUV5 promoter in DE3 lysogen. This expression will in turn cause the expression of the target gene. Such expression is often selectively disadvantageous, and in the present case, clones expressing the recombinant proteins appeared to be overrun by degenerate clones within 25 to 30 generations.

In contrast, dynein cDNA up to 8.7 kb in length was successfully inserted into pET24 vectors and stably maintained in BL21(DE3), and recombinant protein of over 340 kDa could be expressed. Usually, a very high level of expression was obtained with IPTG induction and the recombinant proteins accounted for as much as 30% of the total cellular protein in
some cases. However, the proportion of full-length polypeptides decreased rapidly as the size of recombinant proteins increased, presumably as a result of premature termination of translation or of protease degradation. For example, a full-length 60-80 kDa recombinant protein, such as RB512, accumulated up to 30% of the total cellular protein, while a full-length 300 kDa protein only reached about 1%.

In most *E. coli* expression systems, synthesis of the target protein initiated by the addition of an inducer caused a sudden and overwhelming increase of the nascent polypeptides or folding intermediates. Similar phenomena were also observed in the current study. A high local concentration of folding intermediates may favor aggregation and overload the cellular mechanism that assists protein folding. Reducing the level of expression and changing the way in which the target protein is produced may alter the steady-state concentrations of folding intermediates and shift the folding towards more productive pathways. In clones constructed with pHR and BL21(DE3), the same mechanism that leads to the instability in pET5 expression system was used to drive a low-level constitutive expression of the target protein. The stability of the expression construct was achieved by a further reduction in the expression level by incorporating the low-copy number pA15 replicon in the expression vector, and using chloramphenicol resistance as the selection marker instead of ampicillin. Unlike β-lactamase that confers ampicillin resistance, chloramphenicol acetyltransferase is not secreted.
into the medium. Therefore, chloramphenicol selection is maintained even at high cell density.

Solubility of the recombinant dynein proteins

The formation of insoluble inclusion bodies is a common phenomenon in the expression of cloned genes in *E. coli*, and its mechanism is not clear. It does not seem to be directly related to the size of the expressed polypeptide, for proteins as small as 20 kDa may form insoluble aggregates (Marston, 1986). Rather, it may derive from the aggregation of folding intermediates or misfolded polypeptide chains, especially when they are expressed at high levels (Haase-Pettingell and King, 1988; Brems, 1988; Cleland and Wang, 1990; Goldberg et al., 1991). Inclusion bodies formed from homologous *E. coli* proteins were also observed in *rpoH* mutants which lacks the heat shock response (Gragerov et al., 1991; Kucharczyk et al., 1991).

The formation of inclusion bodies can often be suppressed by lowering the growth temperature (Schein and Noteborn, 1988). Co-expression of molecular chaperons also helps the correct folding of some expressed proteins (Goloubinof et al. 1989; Lee and Olins, 1992).

In the current study, all of the newly expressed recombinant dynein proteins that were not fused to a carrier protein formed inclusion bodies. In general, the formation of inclusion bodies was not affected by lowering the growth and induction temperature to 14-22°C, by co-expressing *E. coli* chaperon 60 (or chaperon 70), or the combined use of both. In five recombinant proteins that underwent in vitro refolding
(AN1480, P748, RB512, 349H and 271H), only one (RB512) was able to form a soluble and stable protein. Fusion to thioredoxin (La Vallie et al., 1993) slightly increased the solubility of small recombinant dynein proteins. However, the solubilization only happened when large quantities of DnaK and DnaJ proteins were co-expressed, and the presence of a small 33-kDa protein that co-purified with the fusion protein appeared to be necessary to maintain its solubility. Co-expression of glutathione-S-transferase fusion protein and DnaKJ proteins would probably have similar results.

The difficulties in obtaining large recombinant dynein proteins may be explained by the complexities involved in the folding of multi-domain polypeptides, and by the lack of proper cellular co-factors that assist folding when these polypeptides are expressed in a heterologous host. In relatively small recombinant fusion proteins and non-fusion proteins, the exposure of hydrophobic amino acid side chains resulting from the truncation of the β-heavy chain is probably a major factor that causes aggregation.

The domain structures of dynein heavy chain

Previous knowledge about domain structure in dynein heavy chains has been mainly derived from studying its proteolytic fragments (Mocz, Tang and Gibbons, 1988; Mocz, Farias and Gibbons, 1991), combined with molecular morphology studies by electron microscopy (Goodenough and Heuser, 1982, 1985; Sale et al., 1985).

In this investigation, both the high solubility of refolded RB512 and its secondary structure composition, which
is typical of an average protein, support the notion that this portion of dynein heavy chain may correspond to a protruding structure on the surface (Gibbons et al., 1991a). Since such a structure will most likely be able to avoid the exposure of hydrophobic interface associated with the truncation of the polypeptide chain. On the other hand, the poor solubility of recombinant proteins corresponding to the nucleotide binding domains suggests these domains are more tightly associated with each other or other part of the heavy chain in the native protein.

The most common way of generating repeated sequences in protein structure is though to be through the duplication of an ancestral gene. Such events were proposed to be the mechanism that lead to the generation of the internally repeated domains in the proteins belonging to the immunoglobulin superfamily (reviewed in Williams and Barclay, 1988). The four consensus P-loop motifs in the dynein heavy chains are a highly conserved feature that is found in all forms of dynein. It is plausible that, in the dynein heavy chain, the four P-loop motifs may have arisen through repeated duplication of a single ancestral gene which encoded the nucleotide binding/hydrolysis domain. One line of evidence supporting this hypothesis is the substantial homology between the P-loop motifs at P1 and P3, and the regions about 100 residues downstream from the P-loop (Eshel et al., 1993). Such duplications would presumably be very ancient, and considerable divergence has accumulated among the P-loops. However, the distribution of stretches of conserved amino acid
sequence of these domains should be consistent with their overall patterns of fold, which should be relatively conserved in sequences that constitute the structural and catalytic region of a domain. Exterior residues tend to be highly degenerate since they function chiefly in maintaining a reasonably polar surface (Bowie et al., 1990). A hinge region joining two domains can also be hypervariable, if it does not specify the interaction between the domains.

Based upon the above hypothesis, the nucleotide binding domains in dynein heavy chain were analyzed through multiple alignment of the sequences from a number of isoforms. The distribution patterns of conserved sequence suggested a rather large NTP binding domain which contained some 180-250 residues (Figure 3.15).

The highly conserved P1 region contains two stretches of unbroken sequences (Figure 4.1). The first section of ~200 residues that includes the P-loop was proposed to form the ATP binding pocket (Mocz, 1995). The second highly conserved section about 100 residues downstream from P1 was proposed as part of the ATP-sensitive microtubule-binding sites in dyneins based on amino acid sequence and possible secondary structure homology with kinesin (Wilkerson, King and Witman, 1994).

The sequences surrounding P2 and P3 are highly conserved in cytoplasmic dyneins, especially in regions downstream from the P-loops (Figure 4.2). However, they are much less conserved in axonemal dyneins. This suggests that P2 and P3 may not have a critical role in the common functions of the
dynein heavy chains, but may be involved in regulating the functions that are specific for each class of the heavy chain, such as differences in substrate specificity, maximum ATP hydrolysis rate and sensitivity to microtubules (Paschel and Vallee, 1987; Shpetner, Paschal and Vallee, 1988). The region around P4 is also highly conserved in cytoplasmic dyneins, and moderately conserved in axonemal dyneins (Figure 4.3), suggesting it may have some structural and functional roles in all forms of dynein.

The homology analysis in the current study also identified a region between 200 to 900 residues downstream from P4 which contains two long stretches of unbroken conserved sequences of ~170 and ~500 residues. Further analysis of this region using the computer program "coil" in PC/GENE, which was written according to the method of Lupas, VanDyke and Stock (1991), revealed three stretches of sequence with high probability of forming α-helical coiled-coil conformation (designated C1, C2 and C3, Figure 4.4) in this region. The location of all three hypothesized coiled-coils are highly conserved in all forms of dynein heavy chains, and the lengths of C1 and C2 are highly conserved too. C1 and C2 overlapped with two putative α-helical coiled-coils that have been identified previously in sea urchin dynein β-heavy chain (Gibbons et al., 1991a), and several other dyneins (Mitchell and Brown 1994), which have been proposed to form the short projection on the globular head (Gibbons, et al., 1991a). Therefore, a larger part of the dynein heavy chain could be involved in the formation of the projection. This region is
thus proposed to be a domain that forms the projection. The overall conservation in size and secondary structure features of this hypothetical projection domain suggests that it may have important non-catalytic structural or regulatory function. Small in-frame deletions in C1 in the outer arm dynein in Chlamydomonas flagella have been shown to cause slow swimming phenotype without changing the sliding velocity of the axoneme, possibly as a result of altered interaction between dynein isoforms (Porter et al., 1994).

C. PROPERTIES OF SOLUBLE RECOMBINANT PROTEINS

Three soluble recombinant proteins (RB512, 349TH and 231TH) were tested for their possible biological activities using assays that have been established for native dynein.

The soluble recombinant proteins corresponding to P1 (349TH) and P2 (231TH) were prone to aggregate after they were purified. This probably reveals that in native dynein, the NTP binding sites interact, either with each other or with some other part of the heavy chain. Truncation of the polypeptide chain causes the exposure of the hydrophobic interfaces that are unable to be shielded completely with the fusion of thioredoxin in the recombinant protein. Such hydrophobic interactions may constitute one of the mechanisms which regulate the ATPase activities or the affinity to microtubules under various physiological conditions. However, the fact that both fusion proteins and a specific host protein can form complexes of consistent composition suggests that the fusion proteins may have a rather defined conformation,
possibly one that is determined by their unique amino acid sequence.

In tests for functional interactions, 349TH, which also contains one of the proposed microtubule binding regions in dynein, slightly inhibited the restoration of beat frequency in KCl-extracted sperm axoneme by 21S dynein, but not the beat frequency of reactivated sperm. Although impurities or co-purified host protein could interfere with this assay, a partial reversal of this inhibition by 21S dynein suggests that the interaction between the fusion protein and the KCl-extracted axoneme may be specific. It is conceivable that if 349TH does bind to microtubules through the proposed microtubule-binding region, it may occupy the microtubule surface at high density, similarly to the kinesin head (Harrison et al., 1993; Song and Mandelkow, 1993). A high ratio of 349TH to 21S dynein may be needed to exhibit significant competition on the rebinding of 21S dynein. A further increase in the effective concentration of the fusion protein would require improving the solubility and the stability of the fusion protein during the assay.

In the microtubule co-sedimentation assay, polymerized microtubules were found to disintegrate upon incubation with the fusion proteins (349TH and 231TH), even in the presence of stabilizing agent taxol (Parness and Horwitz, 1981). This effect could have been caused by the presence of certain contaminating proteins in the fusion protein preparations. Aggregation of the fusion protein at a lower pH (pH 7.5) made
it impossible to determine whether the fusion protein by itself had any direct interaction with microtubules.

Vanadate mediated photo-cleavage under V1 condition probably requires a high degree of structural integrity, since it is inhibited by partial proteolytic digestion of the dynein heavy chain (Mocz et al., 1988). Therefore, it is not surprising that 349TH, which contains the ATP binding and hydrolysis domain, did not cleave under the standard V1 cleavage condition. The photolysis of 349TH observed under millimolar concentration of vanadate and prolonged irradiation is reminiscent of that of the V2 cleavage condition (Tang and Gibbons, 1987) (Figure 3.22). Although without direct evidence, the formation of discrete bands similar in size to that of the hypothesized cleavage products, and the parallel lose of 349TH band intensity without significant smearing, strongly suggest that the cleavage must have happened in a specific manner. However, unlike the V2 cleavage, the cleavage here did not show any preference for Mn"" over Mg"", or ADP over ATP. It is possible that changes in protein conformation as a result of truncation may have significantly deviated the cleavage condition. Improvement in the configuration and the purity of the fusion protein may help to develop a more efficient and specific cleavage condition.

The observed CD spectrum of refolded RB512 resembles those of typical globular proteins (Yang et al., 1986). Its predicted content of α-helix (25%) is identical to that of intact dynein (25%), and higher that of Fragment A (22%) (Mocz and Gibbons, 1990), which is in agreement with a higher helix
composition expected for this region. Its total secondary structure components (82%) is also similar to that of native dynein (76%) or Fragment A (87%). These facts suggest that the refolded RB512 may have a conformation similar to that in native dynein.

The refolding of RB512 provided a line of evidence that the hypothetical projection domain downstream from the P4 domain may constitute an independently folding domain, since RB512 makes up about 70% of this region. Such an independently folding domain would probably be required to form the short projection on the globular head, which has been suggested to form the B-link which tethers the outer arm to the B-subfiber in situ in the axoneme (Goodenough and Heuser, 1984, 1989). Although refolded RB512 did not demonstrate any functional interaction with microtubules or 21S dynein in the current assays, the properties observed in this study make it, and possibly the hypothetical projection domain, ideally suited for future functional and structural studies, such as domain swapping between dynein isoforms or crystallography.
Figure 4.1  Alignment of dynein heavy chain sequences in the P1 region.

Multiple alignment of dynein heavy chain sequences in the region surrounding P1 is presented graphically. The highly conserved regions are represented by I, moderately conserved regions are represented by X, and non-conserved regions are represented by ::. The gaps are represented by '.

The GenBank accession numbers for the dynein heavy chain sequences used for the alignment of axonemal group are: *Tripneustes gratilla*: DYH7a, U03978; DYH7b, U03979; DYH7c, U03980; DYH5a, U03977; DYH5c, U03976; DYH2, X59603; and DYH4, U03973. *Chlamydomonas reinhardtii*: alpha, L26049; beta, U02963. *Paramecium tetraurelia* β-heavy chain, U19464. The sequences used for the alignment of cytoplasmic group are: *Drosophila melanogaster*, L23195; rat brain cytoplasmic, L08505; *Caenorabditis elegans*, L33260; *Dictyostelium discoideum*, Z15124; *Aspergillus nidulans*, U03904; *Neurospora crassa*, L31504; and *T. gratilla* DYH1a, Z21941. The above sequences except DYH4 and *Dictyostelium* cytoplasmic dynein were also used for the mixed alignment of axonemal and cytoplasmic dyneins. Positions of three highly conserved sequences, GPAGTGKT, WGCFDENFL and FITMNPYG, are indicated.
Figure 4.2 Alignment of dynein heavy chain sequences in the P2 and P3 regions

Multiple alignment of dynein heavy chain sequences in the region surrounding P2 (A) and P3 (B). The graphic presentation symbols are identical to those in Figure 4.1. The same sets of dynein heavy chain sequences used in the alignment of P1 were also used here. The conserved P-loop sequences are indicated.
Figure 4.3 Alignment of dynein heavy chain sequences in the P4 region

Multiple alignment of dynein heavy chain sequences in the region surrounding P4. The graphic presentation symbols are identical to those in Figure 4.1. The same sets of dynein heavy chain sequences used in the alignment of P1 were also used here. The conserved P-loop sequences are indicated.
Figure 4.4  Schematic diagram of the hypothetical projection domain

Multiple alignment of dynein heavy chain sequences identified a 700-residue conserved region ~200 to ~900 residues downstream from P4. This region is consisted of two long stretches of unbroken sequences about 170 and 500 residues in length (represented by the open boxes). Short gaps are occasionally encountered in the second section, and the location of two largest one (4 and 6 residues) are indicated by double lines in the open box. The location of three highly conserved α-helical coiled-coil structures, designated C1, C2 and C3, are marked by ‘*’. The length of C3 is more variable among dynein isoforms. The coiled-coil structures correspond to C1 and C2 in sea urchin dynein β-heavy chain were proposed to form the short projection on the globular head, or may involved in protein-protein interaction (Gibbons et al., 1991a). Studies here show that these 700 residues may constitute a single domain that can be fold independently from other part of the heavy chain molecular. The corresponding position of RB512 which was expressed and refolded in the current investigation is marked by double lines. The heavy chain sequences used for this alignment are: cytoplasmic dynein from Drosophila, rat brain, human, C. elegans, Dictyostelium, Aspergillus, Neurospora and S. cerevisiae; axonemal dynein from T. gratillia β-heavy chain, Chlamydomonas β- and γ-chain (U15303) and Paramecium β-chain.
A 13,168 bp single cDNA insert that encodes sea urchin dynein β-heavy chain was molecularly cloned in a plasmid vector. The cloned sequence was free from detectable cloning errors, such as deletion and null mutations. Recombinant heavy chains of 30- to 300-kDa were expressed in a variety of E. coli host/vector systems which cross-reacted with anti-dynein antibodies. Soluble recombinant proteins of medium to small size (30- to 60-kDa) were obtained by refolding purified protein, or by co-expressing chaperon 70 protein. Despite they behaved differently from native dynein in several biological and biochemical assays of functional activity, including interaction with microtubules, or vanadate mediated photocleavage, the recombinant proteins appeared to have gained certain native-like conformation.

The experience gained in this study should benefit future research involving the cloning and expression of RNA messages that encode large polypeptides. The domain structure analysis in regions surrounding the P-loops through sequence homology between dynein isoforms should provide important information to guide future studies on the structure organization of dynein heavy chain. The recombinant protein RB512 deserves further study, since it represents the region that is speculated of forming the projection on the globular head, and is one of the highly conserved and presumably functionally important regions. In addition, it can be produced in a very stable soluble form. The solubility and purity
of recombinant proteins representing nucleotide binding domains can possibly be improved by redesigning the polypeptide to be expressed, using different carrier proteins in the fusion protein, using a different expression system with a controllable promoter, and a more specific affinity purification method. A purer preparation of fusion protein will allow more detailed study of its properties and of the feasibility of using it as a model to study the structure-function relationship in native dynein.
APPENDIX I

SYNTHETIC OLIGONUCLEOTIDES USED IN THIS STUDY

MINX2 5'-TGATGGCAGT AGGTAG(AG) AAGGACT-3'
MRSQ2 5'-AACTCCATGA AGGACTACAA (TC)AA-3'
MPNI3 5'-TTAAAACAAGA ATGCGGCGC AATTTAGACC TGGAGAAGTA AGG-3
PCNI1 5'-TTTTGAATTC AGGCTTTGTA TGAGGGTTGG CTAAGC-3'
MMNX1 5'-CTTGGTTCG ACATCAAAAT AGAC-3'
1721N 5'-AGCGCTAGCC AAGAAGGTCA TGAGAAC-3'
1798N 5'-AGCGCTAGCG ATGATGACAA GCACTGC-3'
2069E 5'-TTTGAATTC TGTCATCTGA GACGAT-3'
2087N 5'-AGCGCTAGCC GTGAAGGGA CATGGA-3'
2317E 5'-TTTGAATTTCT TCTTGAAGCG TACCCT-3'
1798B 5'-AGCGGATCCG ATGATGACAA GCACTGC-3'
TRX\_NHE 5'-CAGCGCTAGC GATAAAAATTA TTCACCTGAC T-3'
TRX1ETK5 5'-GGACTAGTGC TAGCTTTTATC ATCATCGTCA CCTCGGAGC
          TGTCAGGGT AGCGTCAAGA AACTC-3'
T3 5'-ATTAAACCCTC ACTAAAG-3'
T7 5'-AATACGACTC ACTATAAG-3'
SK 5'-TCTAGAACTA GTGGAAC-3'
KS 5'-CGAGGTCGAC GGTATCG-3'
APPENDIX II
ALIGNMENT OF NUCLEOTIDE BINDING DOMAINS IN DYNEIN HEAVY CHAINS

Multiple alignments of the amino acid sequences surround the consensus P-loop motif of dynein heavy chain isoforms were generated by the computer program Clustal with the default parameters (Higgins et al., 1992). The amino acid sequences used in the analysis were: sea urchin dynein β-heavy chain (DYH2) (Gibbons et al., 1991); sea urchin dynein heavy chain 1a, 1b, 5a, 7a and 7b (DYH1A, DYH1B, DYH5A, DYH7A and DYH7B) (Gibbons, B.H. et al., 1994); the α and β heavy chain of Chlamydomonas (A$CHL and B$CHL) (Mitchell and Brown, 1994); Drosophila cytoplasmic dynein (DROME) (Li et al., 1994); rat brain cytoplasmic dynein (C_RAT) (Mikami et al., 1993); Dictyostelium cytoplasmic dynein (DICTY) (Koonce et al., 1992); and yeast cytoplasmic dynein (YEAST) (Eschel et al., 1993). Single letter amino acid code is used. *, asterisk indicates that a position in the alignment is perfectly conserved; ., dot indicates that a position is well conserved. The numbers in the right column indicate amino acid position relative to the start of the aligned sequence.
Alignment in the P2-P3 region:

**Alignment of DICTY, DROME, C RAT, DYH1A, YEAST, DYH1B, A$CHL, DYH2, B$CHL, DYH7A, DYH7B, DYH5A, and DYH5B.**

**Sequence Alignment:**

```
<table>
<thead>
<tr>
<th>DICTY</th>
<th>DROME</th>
<th>C RAT</th>
<th>DYH1A</th>
<th>YEAST</th>
<th>DYH1B</th>
<th>A$CHL</th>
<th>DYH2</th>
<th>B$CHL</th>
<th>DYH7A</th>
<th>DYH7B</th>
<th>DYH5A</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVTMNP---GYAGRSNLPDNLKKLFRSMAMIKPDREMIAQVMLYSQGFKTAEVLA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FITI{NF---GYAGHSNLPDNLKKLFRSLAMTTPDRQLlAEVMLFSQGFRSAEKLA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FITMNP---GYAGRSNLPDNLKKLFRSLAMTKPNQ1QVMLYSQGFRTAEKLA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEAST</td>
<td>FITLN---GYNGRSENLPKSKFSMPSQSGQTAEMIQMGFEDXKLSA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1B</td>
<td>FITLPAGKGYGGRQKLPDNLKHVFRPVAMSRPDEQIAEVLFSFGKDSKL 55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A$CHL</td>
<td>FITMNP---GYIGRAELSALKFRTIHMVPMQDLILCEMLEILVLYSGCFIKARPSL 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH2</td>
<td>FITMNP---GYAGRTELPSALKFRPACMVVPDFELICEMLVAGFLDARLLA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7A</td>
<td>FITMNP---GYAGRSDLPSNLKVLFRTAMVMDYALISIEVLVSICGFIKARPSL 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7B</td>
<td>FITMNP---GYAGRSDLPSNLKVLFRTVMVDPKDIAGELISIEYSMPFVDSLRA 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH5A</td>
<td>FITMNP---GYAGRTELPSNLKFRPAMVMDYALISIEVLVSICGFIKARPSL 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

---

**Sequence Alignment:**

```
<table>
<thead>
<tr>
<th>DICTY</th>
<th>DROME</th>
<th>C RAT</th>
<th>DYH1A</th>
<th>YEAST</th>
<th>QMDQLRKKIQEIAKQRHLVTKQE------WVEKILQLHQLNINHGVMVPSGG 207</th>
</tr>
</thead>
<tbody>
<tr>
<td>DROME</td>
<td>NIDEASVAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C RAT</td>
<td>AVDEGEIAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1A</td>
<td>DVNREGTISENLPFEQILIQSCENTMKPKVAEIDLPLLNLSSLDFFGIGIGIT-----VYQK 159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEAST</td>
<td>-------ISEFPEGHETVSLKRLRLLPLGDSDELVKSIFDSDASAG-----TLP 135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1B</td>
<td>KVQEP------KLVQALRVNLTLKLTLDGFVGDFVDFGEVE---KDI 149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A$CHL</td>
<td>-------VEADVFRLAFRDNPIKTLAQDMIFMGMLNDLFG-PFD-RPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH2</td>
<td>-------PBDVLMRALRDFNPKIVSDTFPMGMLGLDLPF-ALD-VPRR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B$CHL</td>
<td>-------SEDDVLRALEDFNLKTLATDTPSFMGLNLDFKTLFLVPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7A</td>
<td>-------DEDILMLSELDNLKPLESDLPFAGTSFLDDFGV-IPK 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7B</td>
<td>-------HESVLRLALKNLKPLFPLAKDVPFLFGLIDFLPIT-----LTEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH5A</td>
<td>-------GEDIVLRALEDSLNPKFLADAILFGLIDFFGCV-LEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

---

**Sequence Alignment:**

```
<table>
<thead>
<tr>
<th>DICTY</th>
<th>DROME</th>
<th>C RAT</th>
<th>DYH1A</th>
<th>YEAST</th>
<th>RKFASLYYLLEDLLSPQKHYDWGLRAIKSVLVVAGSLKRGDPQR----------- 96</th>
</tr>
</thead>
<tbody>
<tr>
<td>DROME</td>
<td>NIDEASVAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C RAT</td>
<td>AVDEGEIAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1A</td>
<td>DVNREGTISENLPFEQILIQSCENTMKPKVAEIDLPLLNLSSLDFFGIGIGIT-----VYQK 159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEAST</td>
<td>-------ISEFPEGHETVSLKRLRLLPLGDSDELVKSIFDSDASAG-----TLP 135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1B</td>
<td>KVQEP------KLVQALRVNLTLKLTLDGFVGDFVDFGEVE---KDI 149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A$CHL</td>
<td>-------VEADVFRLAFRDNPIKTLAQDMIFMGMLNDLFG-PFD-RPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH2</td>
<td>-------PBDVLMRALRDFNPKIVSDTFPMGMLGLDLPF-ALD-VPRR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B$CHL</td>
<td>-------SEDDVLRALEDFNLKTLATDTPSFMGLNLDFKTLFLVPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7A</td>
<td>-------DEDILMLSELDNLKPLESDLPFAGTSFLDDFGV-IPK 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7B</td>
<td>-------HESVLRLALKNLKPLFPLAKDVPFLFGLIDFLPIT-----LTEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH5A</td>
<td>-------GEDIVLRALEDSLNPKFLADAILFGLIDFFGCV-LEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

---

**Sequence Alignment:**

```
<table>
<thead>
<tr>
<th>DICTY</th>
<th>DROME</th>
<th>C RAT</th>
<th>DYH1A</th>
<th>YEAST</th>
<th>RKFASLYYLLEDLLSPQKHYDWGLRAIKSVLVVAGSLKRGDPQR----------- 96</th>
</tr>
</thead>
<tbody>
<tr>
<td>DROME</td>
<td>NIDEASVAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C RAT</td>
<td>AVDEGEIAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1A</td>
<td>DVNREGTISENLPFEQILIQSCENTMKPKVAEIDLPLLNLSSLDFFGIGIGIT-----VYQK 159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEAST</td>
<td>-------ISEFPEGHETVSLKRLRLLPLGDSDELVKSIFDSDASAG-----TLP 135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1B</td>
<td>KVQEP------KLVQALRVNLTLKLTLDGFVGDFVDFGEVE---KDI 149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A$CHL</td>
<td>-------VEADVFRLAFRDNPIKTLAQDMIFMGMLNDLFG-PFD-RPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH2</td>
<td>-------PBDVLMRALRDFNPKIVSDTFPMGMLGLDLPF-ALD-VPRR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B$CHL</td>
<td>-------SEDDVLRALEDFNLKTLATDTPSFMGLNLDFKTLFLVPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7A</td>
<td>-------DEDILMLSELDNLKPLESDLPFAGTSFLDDFGV-IPK 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7B</td>
<td>-------HESVLRLALKNLKPLFPLAKDVPFLFGLIDFLPIT-----LTEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH5A</td>
<td>-------GEDIVLRALEDSLNPKFLADAILFGLIDFFGCV-LEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

---

**Sequence Alignment:**

```
<table>
<thead>
<tr>
<th>DICTY</th>
<th>DROME</th>
<th>C RAT</th>
<th>DYH1A</th>
<th>YEAST</th>
<th>RKFASLYYLLEDLLSPQKHYDWGLRAIKSVLVVAGSLKRGDPQR----------- 96</th>
</tr>
</thead>
<tbody>
<tr>
<td>DROME</td>
<td>NIDEASVAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C RAT</td>
<td>AVDEGEIAENLPFEQILIQSCENTMKPKVAEIDLLPLGLSLFDFVREFIX---YTRA 160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1A</td>
<td>DVNREGTISENLPFEQILIQSCENTMKPKVAEIDLPLLNLSSLDFFGIGIGIT-----VYQK 159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEAST</td>
<td>-------ISEFPEGHETVSLKRLRLLPLGDSDELVKSIFDSDASAG-----TLP 135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH1B</td>
<td>KVQEP------KLVQALRVNLTLKLTLDGFVGDFVDFGEVE---KDI 149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A$CHL</td>
<td>-------VEADVFRLAFRDNPIKTLAQDMIFMGMLNDLFG-PFD-RPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH2</td>
<td>-------PBDVLMRALRDFNPKIVSDTFPMGMLGLDLPF-ALD-VPRR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B$CHL</td>
<td>-------SEDDVLRALEDFNLKTLATDTPSFMGLNLDFKTLFLVPR 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7A</td>
<td>-------DEDILMLSELDNLKPLESDLPFAGTSFLDDFGV-IPK 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH7B</td>
<td>-------HESVLRLALKNLKPLFPLAKDVPFLFGLIDFLPIT-----LTEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYH5A</td>
<td>-------GEDIVLRALEDSLNPKFLADAILFGLIDFFGCV-LEP 138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
DICTY    IPAGLK-VQKECAAIISQYFEPGGLVHKVLEDAGQR-PHMDFTRLVLVNSFFSL    467
DROME    ----Q-VQDIRALLLFFSSGQTVRTLEYAMDE-HMDFTRLRLSLSSFSM    447
C_RAT    ----Q-GQQATMVPQFTSNGLVTKAHAPKLE-HMDLTRLRLCSLCSFSM    447
DYH1A    ----Q-VQVDSANLVSGFRTSVDGTVIRCLEYAAKQDOHMDFTRLRLSNLSFSM    445
YEAST    ISDSFD-M---ASLTDNIPGTCSNDVLHILG-----VKTLINKLETA    393
DYH1B    FYRALEWQRGQFDFVDTSLV---GVVYGQPGLHRLGEBIFDT-LKIQLRLGQN    414
A$CHL    ISTDYNMIEVLK--KPNLSNKAD---QAM--FEM--YVFVAFIAWFGG--    439
DYH2    VQLCYLSLCLL-TEPTIDACP--KEL---YEL--YVFVAFIAWFGG--    416
B$CHL    AMTIKKILEGIL---PETRTVAGGPPDDKL-LHY--HVFVACFAWFGG--    420
DYH7A    VKLSSLMDLCLFDPDQGTEAIAK-MDEAVVGWIEG--VLYFSVWSWIG--    434
DYH7B    VSLMRLYTLCCDLDEIAAGMGGMNSQIQLTQLGQ--LFLFSVWSWIG--    434
DYH5A    VTTLCCLLSEALL--FQETSNDNLKLETAKLHPLICS---TFVSYLMAIG--    434

DICTY    MNRSHVINVTEYQNLSDFFMPFSQPSNYQTNRTLRLYSMLWGGGMGQLVERENFSK    522
DROME    LDQARRVNLTPNAQHPFPSCADQLEHYFPAKVALVSFLSFAFGDKLVRIDLDG    502
C_RAT    LHQACRNVQAQYHAPFPQMEQELRYQRYLVLAYAILSWLNSDGLSMRRALG    502
DYH1A    VQLAVHLISLYQFQFCN---LDDKSCLVITLIIIKLRLYALAGTSGESRQ-----    443
DYH1B    LPEATRM----------NFAKE---------**VFMTNREQPPDFPRILD    444
A$CHL    -------------------------GLVEKDIQPY---RRNFDFKQFQTTT---    463
DYH2    -------------------------SMFQDVLDY---RVEFSKXITEFTK---    440
B$CHL    -------------------------CMLVDKTVY---RTQFSKXWSWEDK---    444
DYH7A    -------------------------ISGEGSRAKVF---DLLRERLAGLEESESRKFLHIE    468
DYH7B    -------------------------TMNGDSRKKF---DFQFSLISETHDPHPKPSIKIT    468
DYH5A    -------------------------NLIESCDAFAF---DTFIR---NQFDEDNDAK---    459

DICTY    FIQT-IAITPVAN-TPILLIDYVSIDDAN---KSLWKNKVSVESVETKVAS-P    571
DROME    FVRVS-VTVPLPGAAGAPIDYEVENMD---WVPSNKFVIEVETKHAV-P    551
C_RAT    YIRR-ITTPPLPQAPIDYEVENMD---WSPWQAVKFQEVETKHAV-P    551
DYH1A    FRQG-ITTPPLPQAPIDYEVENMD---WVPSNKFVIEVETKHAV-P    549
YEAST    FIQT-INT---YFGHSDQGELSSTYSTIVANDLSFSCSEIFSESVLEAHEVMR    494
DYH1B    YFD---ERTGSLATY---TMQDNLDSADFNN---ERGSLA---PSS---PS-S    474
A$CHL    --------------------------VKEPKGTVKDFYNNPTKQFPWABWTVDIYD---    503
DYH2    --------------------------IKPPNQGTVFDFYIDFEQSKLPSKENVPVFELDP---    480
B$CHL    --------------------------VQFPEKGLVYDFYFQQNCIMVPWEDRVTFQYIP---GD-F    483
DYH7A    FVDPPIKPTMPPKQGVSVDY---FVKEDELGRELWEAEAREAP-PIPKDANA    520
DYH7B    --------------------------SISSGRVBF---DLLRERLAGLEESESRKFLHIE    468
DYH5A    --------------------------TMNGDSRKKF---DFQFSLISETHDPHPKPSIKIT    468

DICTY    DVVIPTVDTRRHDNLHAWLSE--HRPLILCGPPGSGKTMILTSTLRARFPDFEVV    624
DROME    DTVPFPTLTVHRESYLTVLAE---HKPLVLCGPPGSGKTMILTSTLRARLPDEEV    604
C_RAT    DVVPFPTLTVHRESYLTVLAE---HKPLVLCGPPGSGKTMILTSTLRARLPDEEV    604
DYH1A    DVVPFPTLTVHRESYLTVLAE---HKPLVLCGPPGSGKTMILTSTLRARLPDEEV    602
YEAST    DTVPFPTLTVHRESYLTVLAE---HKPLVLCGPPGSGKTMILTSTLRARLPDEEV    602
DYH1B    LVPVTPDVPQVLDFPNWLEADNRPQFIFLPGDPCGCMVLRHCQFRLRSTQIA    529
A$CHL    TVVFVPAETSSLRFFLMMVY---LRKPAFVGGAVKWMTQLVKGKLS-LNEEQI    555
DYH2    AVLVNHETFTRFVFMMDLLE---GRRFVMLVQAVLQGKSVLVDGSTLNXLDGEX    532
B$CHL    SLFVPTVSHLIFTYLTFLSLSV---NKHAYMFVNGTGRSAINUNKLRN-MDITTM    535
DYH7A    SIIVTPDVPQVLDFPNWLEADNRPQFIFLPGDPCGCMVLRHCQFRLRSTQIA    529
DYH7B    DLIIQETARTQIIMFHYAIA---MEVFLLFVGPFTGKSAIANDHLG-MKDRY    565
DYH5A    EMLVPTVDTRRHDNLHAWLSE--HRPLILCGPPGSGKTMILTSTLRARFPDFEVV    550

... ** **
DICTY
AHEALRLFQDRLVETEEKWTDKIDEVALKHFPSVNLDA-ALK------RPILYSNW 882
DROME
AHEALRLFQDRLVVDDSERWRTNENIDLVGOQKHFPGENQEEALQ------RPILYSNW 862
C_RAT
AHEALRLFQDRLVDEERERRSDNIDMVALKHFPSNIDKEKAMS------RPILYSNW 862
DYH1A
AHEALRLFQDRLVDEERERMTDQNIADIAMKHFPSNITKEALG------RPILFSNW 860
YEAST
AYEARNR1PADDRLYGKEKNSPKQELLYETVDPYKLPQDGLNISSS------TSLPLSGL 805
DYH1B
SYEARNR1FCDRLVGVEARNRFDNLNLGTLQADWNAGQILQNLN-------GHYYVSW 791
A$CHL
LHEESERVYADRLVSLDADNYKAAAIKAKYPSVADIDDYKKDKPGPIFCHF 815
DYH2
MHECQRVYGDMDNDQITEAEFLKLEYAKFFEDVDEEAAL------KAKPNICHF 789
B$CHL
VEHCERVFRDINEMADMAKDFERAVTQMDCCGGVAM------EEELPIY---- 789
DYH7A
VHEVFRVYYDRVLDDNDRKWTYVNCMDQSLKE------NFH--TLFAHLD 826
DYH7B
IHEITRYVFYDRLVDDKDRITFEETKETTQSMFKQ------SMD--KVLSSH1T 819
DYH5A
SHBCQRVFHDRLINNEDKKYFNEIMESMKHFPSQ------NIDPETFVTKF 803
*   *   *   *   *

DICTY
LTDYQPVNRSDDLREYVKA-------- 901
DROME
LSKDYPVNRREELREYVA-------- 881
C_RAT
LSKDYPVPQELLRYVKA-------- 881
DYH1A
LSKDYPVPQELLRYVKA-------- 879
YEAST
LSLDYFVNEKDLVNFTBE-------- 824
DYH1B
GART-ETSSGSSLPFGKSLG------ 811
A$CHL
ARGADKAY--DEVADYTSL------ 833
DYH2
ATGIGDPKY--MQVPNWPELNK--I-- 810
B$CHL
ASHASMTYPSVPHTLNSYDDV-- 813
DYH7A
SNSDGVKVEEDDLRTLMFCDFDTPK-- 850
DYH7B
PS--GLKIDDNIRSLFFDFGNPDS 843
DYH5A
IFPGRMKVGAEBADRLYEDIVDLN-- 827
Alignment in the P4 region:

**DICTY**  MVEKGGFWRTSIDHKLKDQIF- -VGACNPPTDAGRVRQLHRPAPILLLDF  53
**DROME**  -VEHGPGYERSDQAWSLERFQ- -VGACNPPTDPRKPLSHFRHRVIIIYDDY  52
C_RAT  -VEHGPGYERSDQAWSLERFQ- -VGACNPPTDPRKPLSHFRHPFVYDDY  52
**YEAST**  M-EKGFQWPTPEKRWWTIERIHI- -VGACNPPTDPRKPMERFTHAAILLYGY  52
B$CHL  V- - - - - - - - KIQKLIEINQMAACMPFT-AGSPFTPMQPQHRFVTPFAQVM
**DYH2**  YEAsT-DRQKLITLEIHKCQYVSQCNQIAGSFINSQRLRHCFVQALSF  44

**DICTY**  PSTSSLTQYIGTFNRAMLKLQP--NLRSFPADNLDAMVEYSESQKRFT-PDIQ  104
**DROME**  PGTESSLQYGTFSRAMLQMRMP-ALRGYAEPILNAMVEYLASQDFRTQDMQ  103
C_RAT  PGPSSSLQYGTFSRAMLQMRMP-LSLRTPEPITAMVEYLASQDFRTQDMQ  103
**YEAST**  PSOSSSLQYGTYEYIKAFLKLV-EFRSTTFPPARASVLYECKAARYSTGQ  103
B$CHL  PNAEITRAYQIDGHFPSSPDV-VAKMSNKVLARCTELHRVNNMINFLPSAVK  94
**DYH2**  PQASQALSTYNSILSQQNLASVSNALQKSLPTVPSTASLHLKQAVQSFPLDTAIAK  99

**DICTY**  AHYISSPELSRDRALAEIQMTDGCSTLEGLVRLWAHEALRLFDQQLDQLVEBEKE  159
**DROME**  PHYVYSPPEPWRVRGICEAIRPLDSLPEGLVRLWAHEALRLFDQQLDQLVEBEKE  158
C_RAT  PHYVYSPPEPWRVRGICEAIRPLDSLPEGLVRLWAHEALRLFDQQLDQLVEBEKE  158
**YEAST**  SHYLLSPPELRLVYRGTATIIGPRQFTLSSRLWYAEWAFPDLRVLGGVEKKN  158
B$CHL  PHYQFNALDSLNSIQTQGLTRAIKEYYREPKV-VALMWHCETRVDRMINEADMA  148
**DYH2**  PHYQFNALDSLNSIQTQGLTRAIKEYYREPKV-VALMWHCETRVDRMINEADMA  153

**DICTY**  WTDKIDEVALKHFPSVN-LD-ALKRPILSWMWTLKQY00V- - - - - - - - - NRSD  202
**DROME**  WTENIDVLQHGPHGND1QSEALQRPIYSLNWLSDYMIPV- - - - - - - - NREE  202
C_RAT  WTEIDMVALKHFPSVND1KPKAMSPIYSLNWLSDYMIPV- - - - - - - - - DQEE  202
**YEAST**  SFQQLYETDVKLPLYQD-LQNISSSTALLEGLSFLDPKFKEV- - - - - - - - - - NQTD  202
B$CHL  KDEFIQVVTKKFPPDCGMVAIEERTPLSYASHASMTYTF-DVFYVANLNYDV  202
**DYH2**  ABEKLVLEAYRKFQFD---VDEEALKKFPIHCHFAPDGYQYKVPNPSVE  202

**DICTY**  LRLEYVKALRKYFYEEELDVPVLFWNEVLWHLRIEDRVRQFQGHALLIGVSQGGK  257
**DROME**  LRLEVHARVLKVCYFEELDVPVLFWNEVLWHLRIEDRVRQFQGHALLIGVSQGGK  257
C_RAT  LRDQYVKALRKYFYEEELDVPVLFWNEVLWHLRIEDRVRQFQGHALLIGVSQGGK  257
**YEAST**  LVSFILTKEFPTCDEEVEFLVPMGVSMDWVRIDLAKQVQGQHMLGASRTGK  257
B$CHL  LRLSTDEKLYNEAANWIDLFQAMTBRVIALIDLFPHNMGLVGGSGK  257
**DYH2**  LNLKILLVEALTDYNEAANVNLVLWPEAMQVCRINRLSEQHGRALLGVSQGGK  257

**DICTY**  SRLSFVAWMMGSLTYYKTVINNNYXSDFDDLRMLKAMRCRCKEYKCIFIFDESN  327
**DROME**  TTLRSFVAWMMGSLTYYKTVINNNYXSDFDDLRMLKAMRCRCKEYKCIFIFDESN  312
C_RAT  TTLRSFVAWMMGSLTYYKTVINNNYXSDFDDLRMLKAMRCRCKEYKCIFIFDESN  312
**YEAST**  TIRTFVSAILKLQYPIKIRHASSLSLDLMLKIKASDCLESRTCLIDSN  312
B$CHL  QSLARLASYICGYEVYQISVSTYGuNDFQENNLQLYKAGTSGFIVTFMTDQ  312
**DYH2**  QSLARLASYISYSLQETDFYQTYLRQGYPDFKLQDATVCMKGLKQKNTVFTMDQAQ  312

**DICTY**  VLESFIFERMTLILLAGGVEGPFQFEGGEEFTLAHMHACTAQRNGNLIDDEEELLYK  367
**DROME**  VLDGQFIFERMTLILLAGGVEGPFQFEGGEEFTLAHMHACTAQRNGNLIDDEEELLYK  367
C_RAT  VLDGQFIFERMTLILLAGGVEGPFQFEGGEEFTLAHMHACTAQRNGNLIDDEEELLYK  367
**YEAST**  LTELAFERMTLILANADIFDFLPQSYSYDKLILNRRNCTRSSLGILQETEELYD  367
B$CHL  IVEBOFLIVYVLDSLSTLGIAIADTPEDKEAIAFNAVNEYAKAGIDSAECWDF  367
**DYH2**  VSYERFIVIMDLAGGSEIFDFQFADDQINESSIGQVNYKVMGQLYTQRENCWKF  367

143
BIBLIOGRAPHY


microscopy, in Cell Movement: The Dynein ATPase, Volume 1. (eds.)


