

Identification of Protein–Protein Interactions of the Major Sperm Protein (MSP) of *Caenorhabditis elegans*

Harold E. Smith and Samuel Ward*

Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721 USA

In nematodes, sperm are amoeboid cells that crawl *via* an extended pseudopod. Unlike those in other crawling cells, this pseudopod contains little or no actin; instead, it utilizes the major sperm protein (MSP). *In vivo* and *in vitro* studies of *Ascaris suum* MSP have demonstrated that motility occurs *via* the regulated assembly and disassembly of MSP filaments. Filaments composed of MSP dimers are thought to provide the motive force. We have employed the yeast two-hybrid system to investigate MSP–MSP interactions and provide insights into the process of MSP filament formation. Fusions of the *Caenorhabditis elegans msp-142* gene to both the *lexA* DNA binding domain (LEXA-MSP) and a transcriptional activation domain (AD-MSP) interact to drive expression of a *lacZ* reporter construct. A library of AD-MSP mutants was generated *via* mutagenic PCR and screened for clones that fail to interact with LEXA-MSP. Single missense mutations were identified and mapped to the crystal structure of *A. suum* MSP. Two classes of mutations predicted from the structure were recovered: changes in residues critical for the overall fold of the protein, and changes in residues in the dimerization interface. Multiple additional mutations were obtained in the two carboxy-terminal β strands, a region not predicted to be involved in protein folding or dimer formation. Size fractionation of bacterially expressed MSPs indicates that mutations in this region do not abolish dimer formation. A number of compensating mutations that restore the interaction also map to this region. The data suggest that the carboxy-terminal β strands are directly involved in interactions required for MSP filament assembly.

© 1998 Academic Press Limited

Keywords: nematode sperm; cytoskeleton; filament assembly; cell motility; immunoglobulin fold

*Corresponding author

Introduction

Nematode sperm have become a useful model system for the study of amoeboid cell movement. Unlike most animal sperm, which are flagellated, these cells crawl *via* an extended pseudopod. This mechanism of motility is observed in a variety of cells, such as human macrophages, and typically utilizes the actin cytoskeleton. However, actin also is involved in additional cellular processes, such as cytokinesis, phagocytosis, trafficking of intracellular and membrane components, and determination of cell shape and growth, which complicate the

characterization of its role in pseudopod formation and crawling. In contrast, the nematode pseudopod contains little or no actin, but instead is filled with filaments composed of the major sperm protein (MSP; Nelson *et al.*, 1982; Roberts *et al.*, 1989). MSP is a small (~14 kDa) basic protein typically encoded by a multigene family of up to 28 members (Scott *et al.*, 1989a). Expression is limited to the sperm, and MSP appears to function solely in cell motility. Thus, nematode sperm offer a simple and appealing alternative for the investigation of amoeboid cell motility.

Light microscopy of crawling sperm from the large pig intestinal parasite *Ascaris suum* reveals a pseudopod filled with MSP filaments organized into mesh-like complexes (Sepsenwol *et al.*, 1989). The complexes appear to treadmill in the direction opposite to the direction of motion, due to the assembly of new MSP filaments at the leading edge of the pseudopod and disassembly at the

Abbreviations used: MSP, major sperm protein; LEXA, *lexA* DNA binding domain; PCR, polymerase chain reaction; FB, fibrous body; MO, membranous organelle; IPTG, isopropyl- β -D-thiogalactopyranoside; Ig, immunoglobulin; ONPG, *o*-nitrophenyl- β -D-galactopyranoside.

trailing edge (Roberts & King, 1991). This pattern of regulated assembly/disassembly resembles that seen for actin filaments in other amoeboid cells. The rate of sperm locomotion is equal to the rate of MSP filament formation, and filament assembly appears to provide the motive force. *In vivo* experiments demonstrate that MSP filament formation is regulated by intracellular pH (King *et al.*, 1994). The leading edge of the pseudopod (pH 6.4) is ~ 0.2 pH unit higher than the trailing edge (pH 6.2). Acidification of the cell results in loss of the pH gradient, rapid disassembly of the MSP cytoskeleton, and disruption of motility; all of these changes are reversed upon removal of the acid.

In vitro analysis of *A. suum* MSP has provided insights into the structure of MSP filaments. MSP exists as stable ($K_d < 5 \times 10^{-8}$) dimers in solution across wide ranges of pH and salt concentrations (Haaf *et al.*, 1996). The MSP cytoskeleton in the *Ascaris* pseudopod is quite labile, but treatment of cells with glutaraldehyde or polyethylene glycol prior to detergent lysis permits recovery of intact filaments (King *et al.*, 1992). Filaments can also be generated by treatment of purified MSP with water-miscible alcohols, and these appear identical to native filaments by electron microscopy. MSP filaments are constructed from two subfilaments that consist of MSP subunits, presumably dimers, arranged in a left-handed helix. The two subfilaments coil around each other to form the right-handed helical filament. The geometry of these structures indicates that the same molecular interactions that join subfilaments into filaments are also available for the association of filaments into complexes (Stewart *et al.*, 1994).

An *in vitro* motility system based on MSP has been reconstituted from *Ascaris* sperm (Italiano *et al.*, 1996). In addition to MSP, the reaction requires membrane vesicles, ATP, and one or more cytosolic components. The membrane vesicles, which are derived from the leading edge of the pseudopodial plasma membrane, nucleate the assembly of MSP into mesh-like fibers, and fiber formation propels the vesicle forward. Vesicle movement resembles the actin-based motility of intracellular pathogens such as *Listeria monocytogenes* and *Shigella flexneri* (Theriot, 1996). The vesicle and cytosolic components required for motility have not been identified, though the presence of a vesicular epitope recognized by anti-phosphotyrosine antibodies correlates well with the ability to promote fiber formation. Also, since MSP contains no nucleotide binding site and fails to bind ATP appreciably, the requirement of ATP for *in vitro* motility must be mediated by one or more non-MSP components.

Recently, the crystal structure of *A. suum* MSP- α has been determined at 2.5 Å resolution (Bullock *et al.*, 1996). The protein forms an immunoglobulin-like fold composed of a seven-stranded β sandwich with a 3_{10} helix between strands *f* and *g*. Its structure is most similar to the bacterial chaperonin PapD (Holmgren & Branden, 1989), although the

amino acid sequences of MSP and PapD share little similarity. The crystal consists of symmetric dimers that associate primarily through the β strand designated a_2 . Additional sites of interaction between MSP dimers were identified in the crystal, but the relationship between these contacts and those involved in filament assembly are unclear. Also, different crystals have been obtained in which MSP appears to form subfilaments similar to those observed *in vivo* (King *et al.*, 1992; Stewart *et al.*, 1993), suggesting that the interaction of MSP dimers within the characterized crystal may not reflect functional interactions required for filament formation.

Mutational analyses of spermatogenesis in the free-living soil nematode *Caenorhabditis elegans* have identified a number of genes required for sperm development (reviewed by L'Hernault, 1997). None of these mutations is found in MSP, presumably because MSP is encoded by a large multigene family. However, several of the genes recovered in these screens are likely to be involved in MSP assembly and motility. In developing spermatocytes, MSP accumulates in a specialized structure, the fibrous body–membranous organelle (FB-MO); the fibrous body dissociates as the cells mature to form spermatids. In *spe-6* (spermatogenesis-defective) loss-of-function mutants, MSP fails to assemble in the FB-MO but instead is distributed throughout the cytoplasm (Varkey *et al.*, 1993). Mutations in the fertilization-defective genes *fer-2*, *-3*, *-4* or *-6* either prevent fibrous body disassembly or promote formation of paracrystalline inclusion bodies composed of MSP (Ward *et al.*, 1981; Ward & Klass, 1982). All of the above mutants are defective in pseudopod formation and amoeboid motility, which demonstrates the importance of regulated MSP assembly in these processes. However, these sperm exhibit additional developmental and/or cytological abnormalities as well, so ascribing a direct role for these genes in MSP filament formation is difficult.

We would like to understand the mechanics and regulation of MSP filament assembly in *C. elegans*. To identify regions of MSP that might be critical for filament formation, we have utilized the yeast two-hybrid system to identify mutations within MSP that abolish interaction. In addition to mutations predicted from the crystal structure, we have identified multiple mutations within the two carboxy-terminal β strands of the protein that prevent MSP–MSP interaction. Biochemical and suppressor analyses suggest that these β strand interactions are likely to be directly involved in MSP filament formation.

Results

MSP–MSP interaction *via* the two-hybrid system

MSP filament formation is thought to generate the motive force in nematode sperm. Mutations

that abolish MSP–MSP interaction should identify residues critical for MSP filament assembly. However, MSP in *C. elegans* is encoded by 28 genes whose products are 97 to 100% identical, and this redundancy precludes a traditional genetic screen for MSP mutants. Therefore, we utilized the yeast two-hybrid system to isolate mutants defective in MSP–MSP interaction. This approach has been used successfully to characterize actin mutations that lie in the predicted filament assembly interface (Amberg *et al.*, 1995). Fourteen of the 28 MSP genes produce identical proteins; we chose one of these, *mSP-142*, for further analysis.

Figure 1A shows a schematic diagram of the two-hybrid system. The *mSP-142* gene was fused in

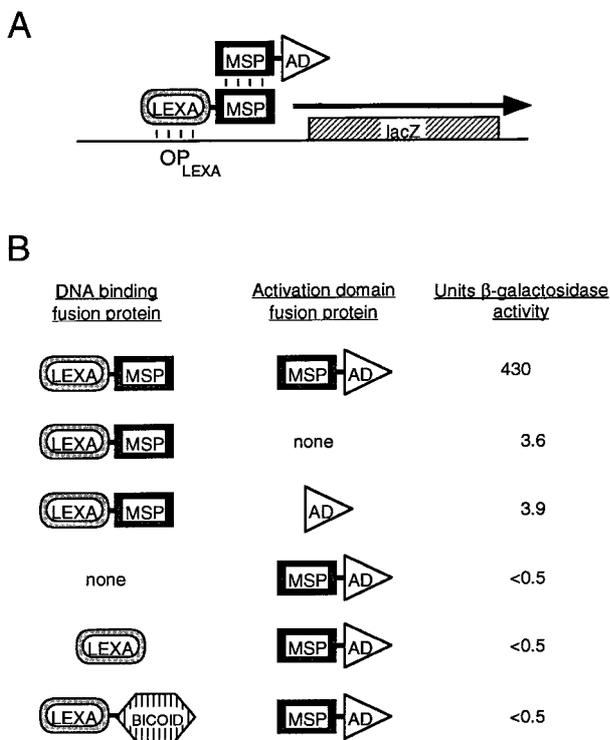


Figure 1. MSP–MSP interaction in the two-hybrid system. **A**, Schematic diagram of the two-hybrid system. MSP is expressed in yeast as protein fusions to a DNA binding domain (LEXA) and to a transcriptional activation domain (AD). *lexA* binding sites (OP_{LEXA}) are located at the promoter region of a *lacZ* reporter gene. LEXA-MSP binds to OP_{LEXA} , but, by itself, is unable to activate transcription of *lacZ*. However, interaction between the MSP portions of LEXA-MSP and AD-MSP recruits the latter protein fusion to the promoter region, where AD stimulates expression of *lacZ*. **B**, Specificity of the MSP–MSP interaction. The indicated combinations of LEXA and AD fusion proteins were co-expressed in yeast that contained the *lacZ* reporter construct pJK103 (Golemis *et al.*, 1997). *lacZ* expression is activated only when MSP is present in both fusion proteins, which demonstrates a specific interaction between MSP molecules. Expression of *lacZ* was determined by assays for β -galactosidase activity (Miller, 1972). Values represent the mean of three or four independent transformants, and variation between samples was <30% of the mean value.

frame to the DNA-binding domain of the bacterial *lexA* gene to create plasmid pLEXA-MSP, and to the transcriptional activation domain B42 (Ma & Ptashne, 1987) to create plasmid pAD-MSP. Both fusion protein constructs were transformed into yeast that contained a *lacZ* reporter gene plasmid whose expression is regulated by *lexA* operators. Interaction between LEXA-MSP and AD-MSP proteins reconstitutes a functional transcriptional activator and drives expression of the *lacZ* gene. Preliminary experiments indicated that constitutive expression of the LEXA-MSP fusion protein was deleterious to yeast; therefore, expression of both LEXA-MSP and AD-MSP was placed under control of the inducible *GAL1* promoter. On glucose (repressing) medium, growth of these MSP-bearing strains was equivalent to non-MSP bearing strains, and no expression of the *lacZ* reporter was observed (data not shown). A shift to galactose (inducing) medium resulted in high levels of *lacZ* expression, as indicated by assays for β -galactosidase activity (Figure 1B). Controls demonstrated that the interaction was mediated by the MSP portions of the fusion proteins: LEXA-MSP, either alone or with AD, induced much lower levels (<1%) of *lacZ* expression. Likewise, AD-MSP alone did not induce detectable *lacZ* expression, and also failed to associate with LEXA or with the heterologous LEXA-BICOID fusion protein. Therefore, the two-hybrid system is able to detect specific MSP–MSP interactions.

Isolation of interaction-defective mutants

To identify mutations that disrupt MSP–MSP interaction, we constructed eight independent libraries of mutant pAD-MSP plasmids (designated pAD-MSP*). The MSP gene was amplified *via* the polymerase chain reaction (PCR) under conditions of reduced fidelity, either by biasing the nucleotide ratios or by adding Mn^{2+} to the reaction (Leung *et al.*, 1989). Mutagenized MSPs were recloned into the activation domain plasmid to generate libraries of pAD-MSP*; this strategy ensured that mutations were limited to the MSP portion of the AD-MSP fusion protein.

Each pAD-MSP* library was introduced into yeast that contained both pLEXA-MSP and the *lacZ* reporter plasmid. Plates containing 2×10^3 to 5×10^3 independent transformants were replica-plated onto SGal-His-Ura-Trp/X-Gal indicator plates (hereafter indicated as X-GAL); this medium maintains plasmid selection, induces pLEXA-MSP and pAD-MSP expression, and detects β -galactosidase activity. The majority of the pAD-MSP* transformants turned blue on X-GAL plates, as did control colonies bearing wild-type pAD-MSP, indicating *lacZ* expression caused by interaction between AD-MSP and LEXA-MSP. On each plate, 0.5 to 1.0% of the transformants were white, representing potential pAD-MSP* interaction-defective mutants. Five of these transformants from each

library (40 total) were selected at random for further analysis.

Each transformant was screened by PCR to ensure the presence of pAD-MSP. PCR amplification was performed with primers specific for the region of pAD flanking the MSP insert. Thirty-six of 40 transformants produced the predicted PCR fragment. The remaining four presumably lack the MSP insert and were discarded. Each pAD-MSP* plasmid was recovered from yeast, transformed into *Escherichia coli*, isolated, and analyzed by restriction mapping. Five of the plasmids could not be recovered, possibly due to plasmid rearrangements and/or deletions (data not shown). Each of the 31 remaining pAD-MSP* plasmids was retransformed into yeast containing pLEXA-MSP and the *lacZ* reporter plasmid, and retested on X-GAL plates for β -galactosidase activity. All produced white or very light blue colonies, compared to dark blue for the wild-type construct.

The DNA sequence of each pAD-MSP* plasmid was determined in order to identify the mutation present. Ten of the mutations were single base-pair deletions within the coding region of MSP that created shifts in the reading frame and would produce grossly aberrant proteins. Two of the mutations were nonsense codons early in the coding region that would generate truncated proteins. One mutation was a conversion of the native stop codon to tryptophan and would create a carboxy-terminal extension of 51 amino acid residues. Twelve of the mutants contained nucleotide changes that would produce single amino acid substitutions. The remaining six plasmids contained two or more mutations; when conveniently located, the individual mutations were separated by subcloning and retested to determine which one yielded the mutant phenotype. Because we were interested in identifying residues critical for MSP–MSP interaction, we chose to focus on the mutations that generate single amino acid substitutions. Fourteen different mutations, two of which were recovered twice, fulfilled this criterion. We also isolated two additional missense mutants in our suppressor screen (described below) that fail to interact with wild-type LEXA-MSP. The 16 different missense mutations that cause defects in MSP–MSP interaction are shown in Figure 2A.

Close examination of the mutations reveals that the distribution is non-random. Of the 18 independent mutants, only 12 of 127 total residues were altered. Two mutations (I123N and Y125C) were recovered twice from different PCR libraries and thus represent *bona fide* independent isolates. Furthermore, multiple mutations were recovered in four amino acids: asparagine at position 36 (N36S, N36I), isoleucine at 123 (I123N twice, I123T), tyrosine at 125 (Y125C twice, Y125H), and asparagine at 126 (N126D, N126K). Unlike most mutagenic agents, amplification by error-prone PCR does not constrain the types of base-pair substitutions (i.e. transitions or transversions) that one can recover, and each nucleotide is a potential target for

mutation, so the process of mutagenesis cannot explain the non-random distribution. Thus, although this screen for mutants was not saturating, it is clear that only a limited subset of residues can be altered that abolish MSP–MSP interaction.

Assays for β -galactosidase activity demonstrate that all of the pAD-MSP* mutants exhibit reduced levels compared to wild-type (Figure 2B). Ten of the 16 mutations reduced activity roughly 100-fold to levels indistinguishable from the pAD plasmid lacking MSP, and thus appear to abolish interaction completely. One of the mutations (Y125C) caused a 30-fold reduction in activity, one (D88V) a 15-fold reduction, and four (F20S, V92A, K119E, and Y125H) caused eight- to tenfold reductions. These six mutations appear to produce proteins that retain some ability to interact with wild-type LEXA-MSP.

Recently, the crystal structure of the *Ascaris suum* α -MSP was determined (Bullock *et al.*, 1996). The protein crystallizes as a dimer, and each monomer forms a seven-stranded β sheet with an immunoglobulin (Ig)-like fold. Given the high degree of similarity (83% amino acid identity) between the *A. suum* α -MSP and the *C. elegans* *msp-142* gene product, it is probable that the two proteins share virtually identical structures. The mutations were mapped onto the crystal structure to visualize their locations within the protein (Figure 3). A schematic diagram of the structure is also shown above the amino acid sequence in Figure 2A.

The interaction defect caused by some MSP mutations can be interpreted in light of their locations in the MSP structure. These mutations comprise two separate classes based on their predicted structural defects. Mutations of the first class (N36S, N36I, and L63R; shown in green in Figure 3) are in residues critical for the overall fold of the Ig-like domain, specifically those proteins similar to PapD (Bork *et al.*, 1994; Holmgren *et al.*, 1992). Mutations in these amino acids would be likely to disrupt the global conformation of MSP. Mutations of the second class are in residues that constitute the dimer interface between MSP monomers, and map to either β strand a_2 (K17E, F20S, and N21Y; shown in red) or residue 126 of strand g (N126D and N126K; also shown in red). These mutations are likely to prevent dimerization.

All of the remaining mutations lie within the carboxy-terminal third of the protein in either strand f (D88V and V92A; shown in purple) or strand g (K119E, L121P, I123N [2x], I123T, Y125C [2x], and Y125H; shown in blue). Neither of the f strand residues has been implicated in protein folding or dimerization, and, though the g strand lies adjacent to the a_2 dimer interface, only the side-chain of N126 is thought to play a role in dimer formation (Bullock *et al.*, 1996). The region of the g strand defined by these mutants (residues 119 to 126) also appears to be distinct from a region proposed from the structure of *A. suum* α -MSP to be involved in filament formation (residues 113 to 120 in

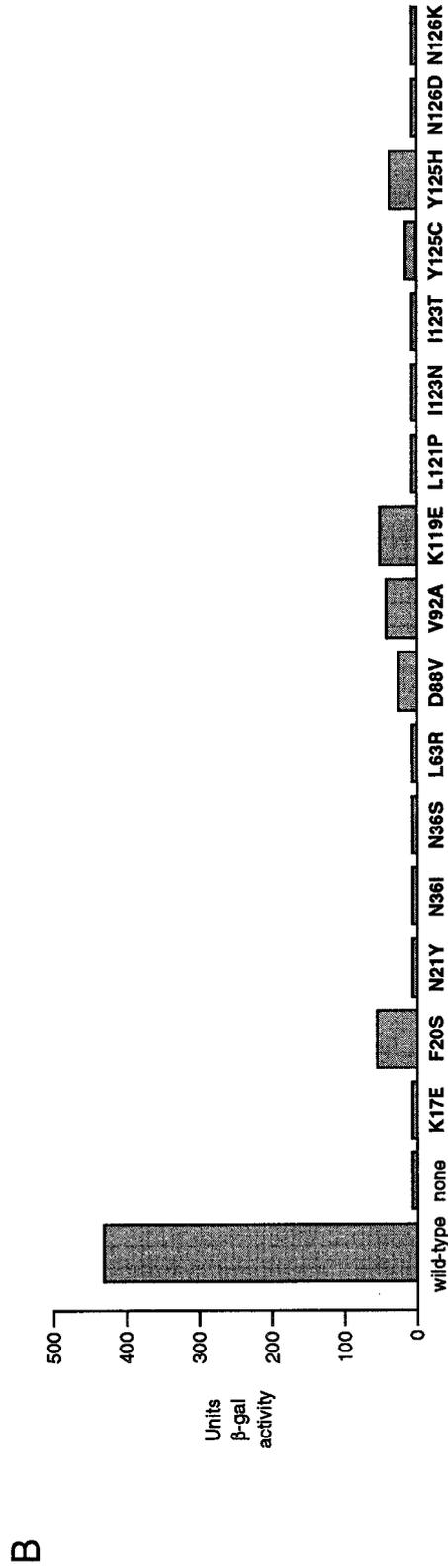
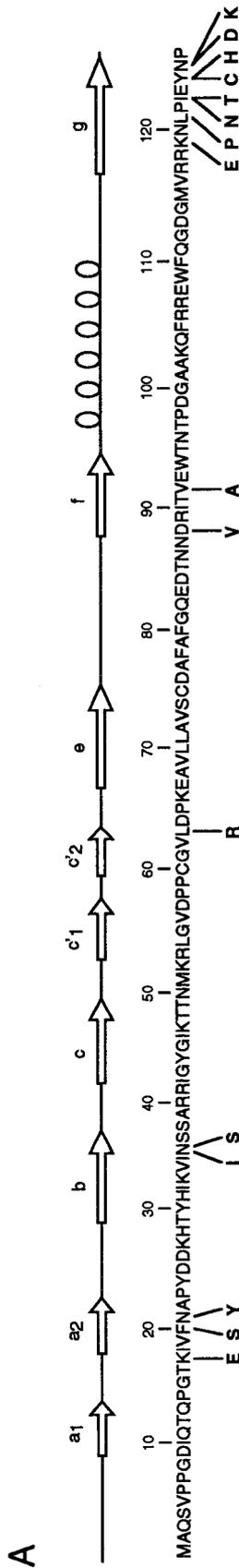
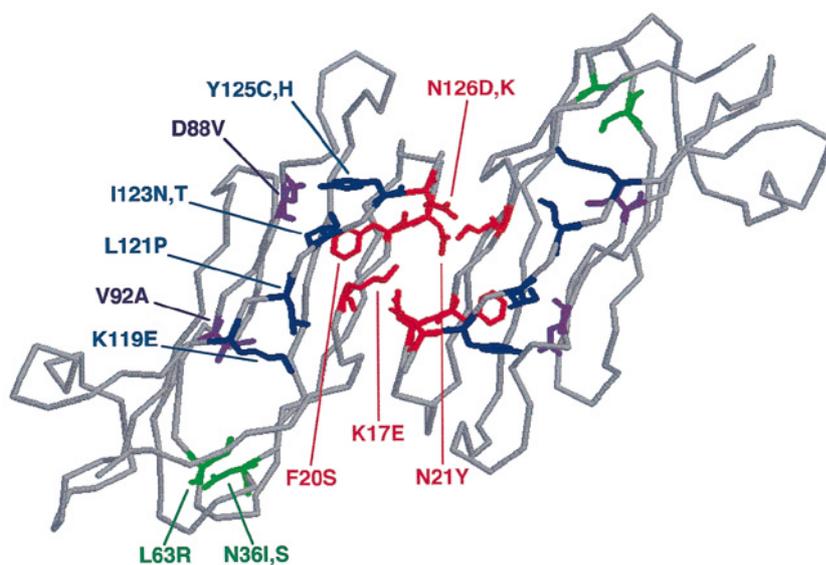


Figure 2. Missense mutations that cause defects in MSP–MSP interaction. **A**, Positions of mutations within MSP. The predicted protein sequence of the *msp-142* gene of *C. elegans* is shown using the standard single letter nomenclature for amino acids. Amino acid substitutions that reduce or eliminate MSP–MSP interaction are indicated below the sequence. The predicted secondary structure based on *A. suum* MSP- α is shown above the sequence. Arrows denote the seven β strands, conventionally labeled *a* to *g*, including the half-strands *a*₁, *a*₂, *c*₁, and *c*₂. Loops denote the 3₁₀ helix between strands *f* and *g*. **B**, Effect of mutations on *lacZ* expression. Levels of β -galactosidase activity were determined for yeast that expressed both wild-type LEXA-MSP and the indicated AD-MSP(*) protein fusion. Mutations are denoted by single letter nomenclature of the wild-type amino acid, its position in the polypeptide chain, and the substituted amino acid; thus, the mutation that replaces the lysine residue at position 17 with glutamic acid is indicated as K17E. Values represent the mean of three or four independent transformants, and variation between samples was <30% of the mean value.



defects resulting from mutation. Those in red (strand a_2 and N126) are defective in dimer formation; those in green (N36 and L63) are defective in protein folding; those in purple (strand f) and blue (strand g) are proposed to be involved in filament assembly. The Figure was generated with coordinates from the Brookhaven Protein Database (identification code 1MSP) with the program RasMol (Sayle, 1996).

C. elegans), as only mutant K119E lies within the latter domain. Because these mutations in the f and g strands seem to identify a novel region for MSP–MSP interaction, we sought to analyze them in greater detail.

One trivial explanation for the apparent interaction defect is that mutations within this region make the protein unstable. The reduction in the level of AD-MSP* protein would manifest itself as a reduction in the level of *lacZ* expression. However, Western blot analysis with antibodies specific to an epitope of AD demonstrated that mutant AD-MSP* protein levels were comparable to wild-type levels (data not shown). Thus, the mutations do not destabilize the protein but must cause defects in MSP–MSP interaction.

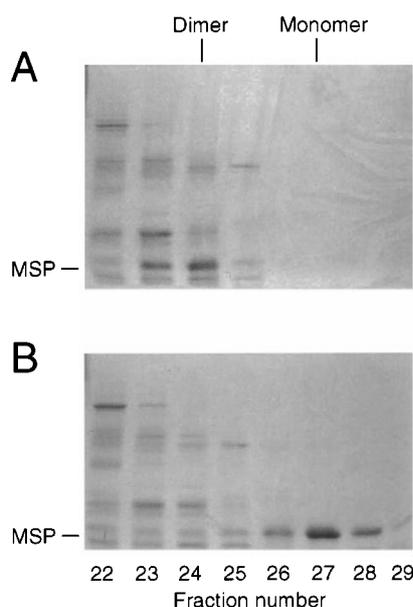
Bacterial expression and dimer determination

MSP molecules strongly associate as dimers in solution (Haaf *et al.*, 1996), and dimers are thought to be the basic building block of MSP filaments. Given the close proximity of the g strand mutations to the dimerization domain (strand a_2), it is possible that the mutations abolish dimer formation. To test this hypothesis directly, and to eliminate any effects of the activation domain, we expressed several of the MSP proteins (wild-type or mutant) in bacteria and determined their dimerization state *via* gel filtration chromatography. Plasmids (designated pET-MSP) were constructed by excising the entire *msp-142* coding region from the appropriate pAD-MSP(*) plasmid and cloning into a bacterial expression vector; this cloning strategy permits IPTG-inducible synthesis of MSP (see Materials and Methods for further details). Pilot experiments indicated that several of the mutant

Figure 3. Position of mutations within the predicted structure of the MSP dimer. Missense mutations recovered in the screen for interaction-defective MSPs were mapped onto the structure of the *A. suum* MSP- α dimer. The C α polypeptide backbone is shown in gray. The side-chain of the wild-type amino acid is shown at the position of each mutation. These positions are indicated in both subunits of the MSP dimer; the mutations are labeled only in the MSP monomer on the left. Amino acid substitutions are denoted as in Figure 2. Numbering of amino acid position is based on the predicted sequence of the *C. elegans* *msp-142* gene product, and is +1 compared to *A. suum* MSP- α . Residues are color-coded by the specific interaction

MSPs were insoluble or unstable under standard growth conditions (i.e. 37°C in LB medium); therefore, MSP expression was induced under conditions shown to improve the recovery of soluble protein (Blackwell & Horgan, 1991; see also Materials and Methods). Despite these efforts, protein from mutant N126D could not be detected; either the plasmid is unstable or the protein is highly toxic in *E. coli*, since deletion derivatives of the plasmid were frequently recovered. Two other mutants, N36S and L63R, produced large amounts of protein that remained in the pellet (i.e. insoluble) fraction after lysis; therefore, these mutations reduce protein solubility. The remaining mutants tested produced soluble MSP as the predominant band on Coomassie-stained SDS/polyacrylamide gels; Western blotting with MSP-specific antibodies confirmed that the band was indeed MSP (data not shown).

Wild-type and mutant MSPs were analyzed by gel filtration chromatography and SDS/PAGE of column fractions followed by Coomassie staining. The results are shown in Figure 4. Wild-type MSP eluted from the sizing column at an apparent relative molecular mass of 28×10^3 (Figure 4A), which is in good agreement with the predicted size of the MSP dimer ($M_r = 28.4 \times 10^3$). Wild-type MSP exhibited the same elution profile when loaded as total protein extract or after substantial (>95%) purification *via* ion-exchange chromatography (Haaf *et al.*, 1996; data not shown). Protein from four of the mutants (K17E, F20S, N21Y, and N126K) eluted from the column with a peak at 14 kDa (see Figure 4B for example), the molecular mass of the MSP monomer. Thus, as predicted by the crystal structure, these mutations abolish dimer formation. Protein from four other mutants



C

MSP tested	Status
wild-type	Dimer
N36S	Insoluble
L63R	Insoluble
K17E	Monomer
F20S	Monomer
N21Y	Monomer
N126K	Monomer
K119E	Dimer
L121P	Dimer
I123N	Dimer
Y125H	Dimer
Y125C	Mixed

Figure 4. Monomer/dimer determination of MSP protein. Wild-type or mutant MSP was expressed in *E. coli* and the protein was size-fractionated by analytical gel filtration. Fractions were collected, concentrated, and subjected to denaturing polyacrylamide gel electrophoresis. Examples of Coomassie-stained gels are shown in A and B. The predicted elution peaks for MSP dimers and monomers are indicated above the gels. The band identified as MSP protein is indicated on the left. The column fraction number is indicated below. A, Elution profile of wild-type MSP. The protein peak elutes at the predicted molecular mass of an MSP dimer (~28 kDa). B, Elution profile of MSP mutant N126K. The protein peak elutes at the predicted molecular mass of the MSP mono-

mer (~14 kDa). C, Summary of results from the various MSP proteins tested. Insoluble, indicates that the protein remained in the pellet following bacterial lysis; Mixed, indicates elution peaks of both dimers and monomers.

(K119E, L121P, I123N, and Y125H) eluted in a peak indistinguishable from wild-type MSP, indicating that these mutations have no effect on dimer formation. Protein from the last mutant tested (Y125C) eluted in two peaks, one dimeric and the other monomeric in size. The presence of monomers might be due to the formation of inappropriate intramolecular disulfide bonds and the resultant protein misfolding, or to an effect on the adjacent N126 residue, which is known to be involved in dimerization. Taken together, the results indicate that mutations in *g* strand residues 119 to 125 do not abolish dimer formation. Therefore, they must disrupt other MSP–MSP interaction(s) that are likely to be involved in either the assembly of dimers into subfilaments or the association of subfilaments into filaments and more complex structures.

Isolation of suppressor mutants that restore interaction

If the mutations within the *g* strand cause a defect in a specific interaction, then it might be possible to obtain compensatory mutations in another MSP molecule that restore the interaction. To permit screening of the existing pAD-MSP* mutant libraries for possible compensatory mutations, six of the *g* strand mutations (K119E, I123N, Y125C, Y125H, N126D, and N126K) were cloned into the pLEXA vector to create pLEXA-MSP* mutant plasmids. When tested in yeast that contained wild-type pAD-MSP and the *lacZ* reporter plasmid, each pLEXA-MSP* mutant was defective in interaction in the two-hybrid system and remained white or very light blue on X-GAL indicator plates. The pAD-MSP* libraries were pooled and then introduced into yeast that contained a

pLEXA-MSP* mutant and the *lacZ* reporter plasmid. Plates with 3×10^3 to 5×10^3 transformants were replica-plated onto X-GAL indicator plates and examined after three days of growth. On each plate, 0.1 to 0.5% of the transformants produced blue colonies, indicative of possible suppressor mutations that restore interaction. Four to eight of these transformants from each plate were selected at random for further analysis.

Each pAD-MSP* plasmid was recovered from yeast, transformed into *E. coli*, isolated, and sequenced. Several of the original interaction-defective mutations were reisolated as suppressor mutations, so we tested each of the 16 pAD-MSP* interaction-defective mutants for the ability to restore interaction with the pLEXA-MSP* *g* strand mutants. Also, identical mutations were recovered as suppressors of different pLEXA-MSP* mutants a number of times; therefore, we tested each pLEXA-MSP* *g* strand mutant and each pAD-MSP* suppressor mutant in every pairwise combination. Assays of β -galactosidase activity were used to quantify the relative levels of interaction between mutants. The results are shown in Figure 5.

The suppressor mutants that restore the interaction with pLEXA-MSP* *g* strand mutants have several distinctive features. First, 13 of the 16 mutants that fail to interact with wild-type MSP function as ubiquitous suppressors; though the levels of β -galactosidase activity vary, each of these mutants restores the interaction with every *g* strand mutant tested. The suppressors include both dimerization-deficient as well as dimerization-proficient mutants. Only those mutations that lie in residues critical for folding (N36I, N36S, L63R) fail to suppress the interaction defect (data not shown). Second, all but one of the remaining suppressors that were isolated contained nonsense mutations

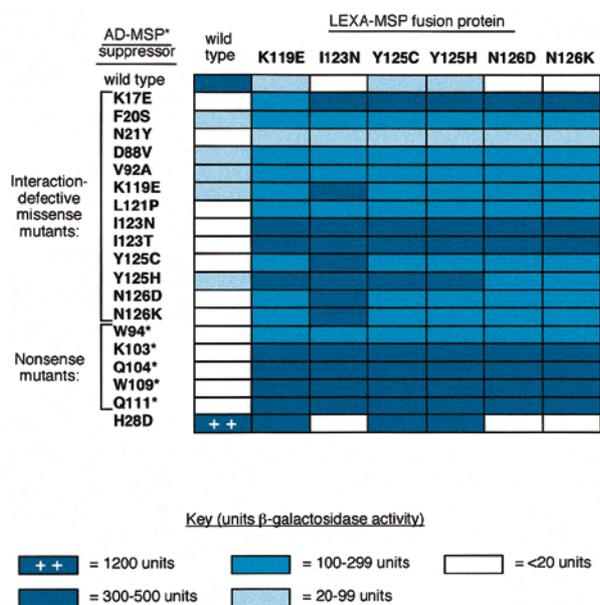


Figure 5. Suppression of interaction defects by MSP mutations. The indicated wild-type or mutant LEXA-MSP(*) was expressed in yeast with the indicated AD-MSP(*) protein fusion. Expression of *lacZ* from the reporter construct pJK103 was assayed as in Figures 1 and 2B. Values represent the mean of three independent transformants, and variation between samples was <25% of the mean value.

(W94*, K103*, Q104*, W109*, Q111*). These nonsense mutations fall within a small range (residues 94 to 111), and share the feature that each retains the *f* strand while removing the *g* strand. Suppression appears to require the *f* strand, because mutations that truncate the protein prior to this region fail to restore the interaction (data not shown). Third, all of the above suppressor mutants interact only weakly or not at all with wild-type MSP. In fact, the interaction-defective mutations D88V and V92A were isolated during the suppressor screen. Finally, a single suppressor was recovered that encodes a novel missense mutation (H28D). This mutant appears to exhibit increased affinity for wild-type LEXA-MSP, as β-galactosidase levels for AD-MSP*(H28D) were threefold higher than for wild-type AD-MSP. Suppression by the H28D mutation is also allele-specific; LEXA-MSP* mutants that interact weakly with wild-type AD-MSP (i.e. K119E, Y125C, Y125H) interact strongly with the H28D mutant, while mutants that fail to interact with wild-type AD-MSP (i.e. I123N, N126D, N126K) also fail to interact with this mutant.

Sequence alignment of known nematode MSPs

MSP proteins are highly conserved among nematode species, and MSP filament formation presumably plays a critical role in sperm motility in these organisms. Therefore, it is likely that the

sites of MSP interactions will lie within the most highly conserved portions of the proteins. Figure 6 shows the alignment of the 18 complete MSP protein sequences currently available. These include ten different MSPs (encoded by 28 genes) from the free-living *C. elegans*, two from the pig parasite *A. suum*, one from the sheep parasite *Dictyocaulus viviparus*, two from the human filarial parasite *Onchocerca volvulus*, and three from the plant parasite *Globodera rostochiesis*. A consensus sequence of amino acids found in all MSP proteins is shown above the alignment.

The alignment indicates that the MSP proteins share a high degree of sequence similarity; the proteins are 61 to 99% identical at the amino acid level. All are likely to exhibit the same structural architecture, since residues known to be important for this type of immunoglobulin-like fold, as well as proline residues that delimit the β strands, are highly conserved (Bork *et al.*, 1994). Three regions of the MSP proteins are extremely conserved: residues 17 to 25, which contain strand *a*₂; 88 to 102, which contain strand *f* and a part of the 3₁₀ helix; and 108 to 126, which contain strand *g*. The *a*₂ strand has been shown to be involved in dimer formation (see above), so sequence conservation probably reflects functional conservation of the dimerization domain. Similarly, conservation of strands *f* and *g* suggests a possible role in MSP filament assembly. Strikingly, all of the interaction-defective mutations, with the exception of those involved in folding (N36I, N36S, L63R), map to one of the three conserved regions, further supporting a functional role for these portions of MSP.

Discussion

We have undertaken an analysis of MSP–MSP interaction to better understand the mechanics of MSP filament assembly and its role in nematode sperm motility. Using a combination of the yeast two-hybrid system and direct biochemical assay of MSP dimerization, we have identified three classes of mutations that map to distinct regions of MSP. Mutants of the first class are defective in protein folding, while mutants of the second class are defective in MSP dimerization. Mutations from both of these classes map to residues predicted by the protein structure. Mutants of the third class are able to form dimers but still fail to interact with wild-type MSP *via* the two-hybrid system. These mutations, which fall within β strands *f* and *g*, define a new site of MSP–MSP interaction that is probably involved in filament formation. Suppressor mutations that restore interaction also fall predominantly within this region. Finally, sequence comparison of known MSP genes identifies three highly conserved regions. Since mutations that disrupt MSP–MSP interactions map to these regions, they probably define functionally conserved domains required for filament assembly.

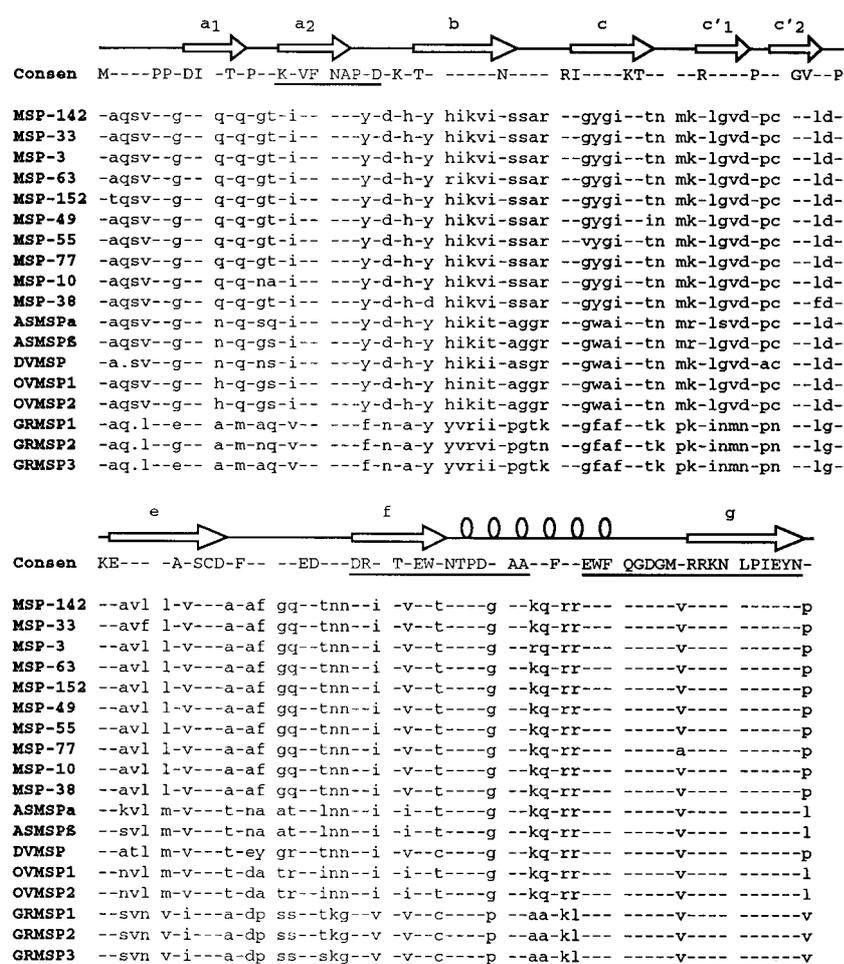


Figure 6. Alignment of known nematode MSPs. The predicted amino acid sequences of all known nematode MSPs are shown using standard single letter nomenclature. A consensus sequence of 100% identical residues is shown in upper case above the alignment. Within individual sequences, the consensus sequence is indicated by dashes; amino acid differences are indicated by lower case letters; and gaps are indicated by dots. Three highly conserved regions of the consensus sequence are underlined. The secondary structure of *A. suum* MSP- α is shown above the consensus sequence. The first ten sequences are from the indicated *C. elegans* gene. Additional sequences are from *A. suum* (ASMSp- α and - β), *D. viviparus* (DVMSp), *O. volvulus* (OVMSp-1 and -2), and *G. rostochiesis* (GRMSp-1, -2, and -3).

Ample precedent exists for use of the two-hybrid system to map regions or residues required for protein–protein interaction. The example most relevant to this study is the analysis of actin–actin interactions *via* the two-hybrid system (Amberg *et al.*, 1995). In most amoeboid cells, actin performs the role that MSP does in nematode sperm: the assembly of filaments necessary for pseudopod formation and cell motility. Mutations that specifically disrupt actin–actin interaction in the two-hybrid system were mapped onto the crystal structure. The mutations correspond to regions of the protein that, based on the F-actin model, are sites of intermolecular interaction between actin monomers (Lorenz *et al.*, 1993). These mutations also define residues critical for the normal function(s) of actin, as all of the mutants exhibit either cold or temperature sensitivity, or lethality, in yeast. Thus, the two-hybrid system is able to identify functionally significant interactions between actin molecules.

Our own two-hybrid results with MSP further validate this approach for isolating relevant mutations. The crystal structure (Figure 3) suggested two potential classes of mutations that might be recovered. One class is predicted to prevent proper protein folding based on the identification of MSP as an Ig-like protein and its close

structural similarity to the bacterial chaperonin PapD. Though the amino acid sequences of MSP and PapD are only 11% identical, four of the 14 conserved residues appear critical for this type of Ig-like fold (Bullock *et al.*, 1996). These four residues (L-D-P at positions 63 to 65 in MSP, 59 to 61 in PapD; N at position 36 in MSP, 24 in PapD) form the identical structure in each protein, with the proline packed against the asparagine side-chain. The interaction-defective mutations in these conserved residues (N36I, N36S, and L63R) are predicted to disrupt folding and alter the global conformation of the protein. Expression of MSP mutants N36S and L63R in *E. coli* produced exclusively insoluble protein (Figure 4); this result is consistent with a folding defect, since misfolded protein would probably form insoluble aggregates. These data confirm the structural significance of these residues in proper protein folding.

The second class of mutations are expected to exhibit defects in MSP dimerization. The a_2 strand (residues 17 to 22) and the asparagine residue at position 126 form much of the interface at the non-crystallographic 2-fold axis and appear to define the MSP dimerization domain. The interaction-defective mutations in these residues (K17E, F20S, N21Y, N126D, and N126K) are predicted to abolish

dimer formation. Expression of mutant MSPs in *E. coli* and subsequent size fractionation confirmed this prediction (Figure 4); mutants K17E, F20S, N21Y, and N126K all produced soluble protein that failed to form dimers. Thus, for both classes of mutants, the analysis of bacterially expressed MSPs not only explains the two-hybrid interaction defects but also provides direct biochemical evidence that the crystal structure of MSP accurately predicts its solution behavior.

The third class of mutations all fall within the carboxy-terminal β strands *f* and *g*. These strands appear to identify a new site of MSP–MSP interaction. None of the altered amino acids lies within the proposed dimer interface. The crystal structure indicates that strand *g*, though not *f*, lies near strand a_2 and might conceivably disrupt dimerization; however, size fractionation indicated that *g* strand mutations K119E, L121P, I123N and Y125H have no effect on dimer formation. All of the mutations are localized to a small portion of the protein; residues 119 to 125 in strand *g* are directly adjacent to the *f* strand mutations D88V and V92A (see Figure 3). These strands are two of the three most conserved regions of the protein based on alignment of known nematode MSP amino acid sequences (Figure 6). (The third region contains the a_2 strand, which is involved in dimerization; see above.) As the only known role for MSP is filament formation in the sperm pseudopod, we believe that the mutations define a site of MSP–MSP interaction involved in this process.

The mutations recovered in this screen appear to identify two structurally and functionally distinct regions of MSP–MSP interaction: one (strand a_2 and residue 126) required for dimer formation and the other (portions of strands *f* and *g*) presumably involved in filament assembly. The residues altered in these two regions also differ in respect to orientation of the side-chain (Figure 3). Three of the four residues involved in dimerization (K17, N21, N126) extend away from the surface of the protein, and appear to form direct contacts between MSP monomers. In contrast, four of the six residues implicated in filament formation (D88, V92, L121, I123) are oriented toward the interior of the protein, and the remaining two (K119, Y125) lie parallel to the surface. Given the relatively inaccessible nature of the amino acid side-chains, a large conformational change would be required to expose these residues for protein–protein contact. Therefore, a direct role in intermolecular association seems unlikely. Rather, we favor the interpretation that the mutations alter the geometry of the *f* and/or *g* strands relative to the core of the protein, and this altered conformation impairs association with wild-type MSP molecules.

This interpretation could explain two surprising results: (1) the recovery of interaction-defective mutants as suppressors that restore interaction; and (2) the lack of allele specificity for all suppressors except H28D (Figure 5). If mutations in *f* and *g* alter only strand topology but not the residues

required for MSP–MSP contact, then interaction mediated by these strands could still occur. Furthermore, suppression by the nonsense alleles suggests that the *f* strand is the site of interaction. Nonsense mutations that remove both *f* and *g* fail to suppress, while those that retain *f* but remove *g* (W94*, K103*, Q104*, W109*, and Q111*) all act as suppressors. Thus, suppression by the *f* or *g* missense mutants as well as the truncation mutants probably occurs *via* the same mechanism, i.e. by promoting interaction between *f* strands. Of course, other interpretations are possible; for example, all of the suppressor mutations could unmask a novel site of interaction that is not accessible in wild-type MSP.

The mutant H28D, isolated in the suppressor screen (Figure 5), differs from the other suppressors. All of the other alleles are defective for interaction with wild-type MSP; in contrast, interaction between wild-type LEXA–MSP and mutant AD–MSP*(H28D) produces threefold higher levels of β -galactosidase activity than wild-type AD–MSP. Mutant H28D is also the only suppressor that is allele-specific; only those mutants that interact weakly with wild-type MSP (K119E, Y125C, and Y125H) are suppressed, and β -galactosidase activity also is elevated relative to wild-type MSP. Thus, H28D appears to interact in the same manner as wild-type MSP but with increased affinity. The crystal structure of MSP indicates that the residue lies adjacent to the dimerization domain a_2 on the side opposite strands *f* and *g*. We speculate that this histidine residue might be critical for the regulation of *in vivo* MSP filament formation by pH. The range of pH that affects MSP assembly and disassembly in the pseudopod is ~ 6.2 to 6.4 , and the imidazole group of histidine is the only side-chain whose charge is significantly affected within this range. Because MSP contains only two histidine residues (the other is located at position 31), these residues may mediate pH regulation of assembly. However, the data are not entirely consistent with this proposal. The substitution of aspartate for histidine would be predicted to mimic MSP under acidic conditions, in which disassembly is promoted, yet this mutation appears to strengthen MSP–MSP interaction.

Model for MSP–MSP interaction

A schematic diagram of our model for MSP–MSP interaction in the two-hybrid system is shown in Figure 7. We propose that monomers of wild-type MSP associate *via* the dimerization domain (strand a_2 and residue 126), and that dimers of MSP associate *via* the filament assembly domain (strand *f*) to form tetramers. Transcriptional activation of *lacZ* expression reflects the formation of either: (1) a heterodimer of LEXA–MSP + AD–MSP, or (2) a tetramer composed of a LEXA–MSP homodimer + an AD–MSP homodimer (Figure 7A). Other combinations are also possible, but are not included for clarity.

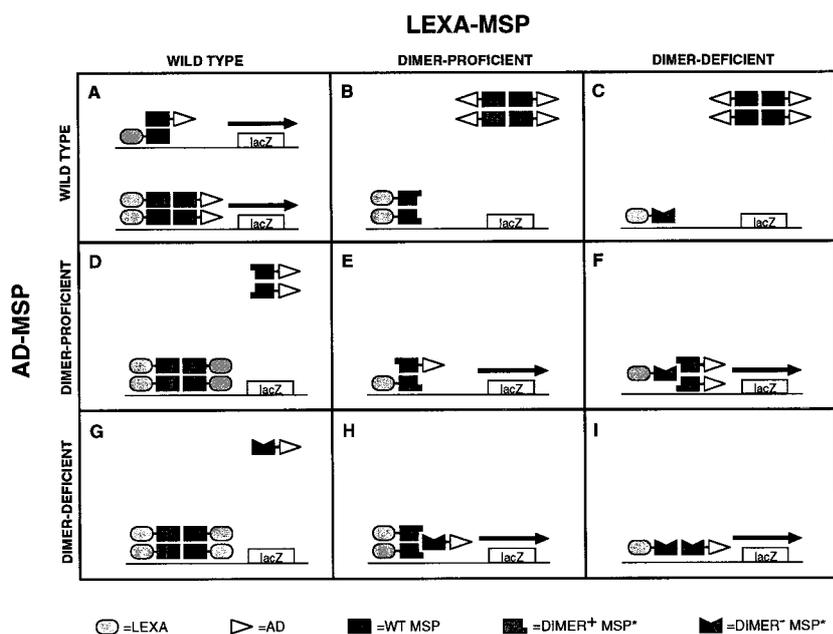


Figure 7. Model figure for MSP–MSP interactions in the two-hybrid system. Symbols are identical to those in Figure 1. Lightly shaded ovals represent LEXA, open triangles represent AD, and darkly shaded boxes represent MSP. An arrow above the *lacZ* reporter gene indicates transcription, and requires interaction between the LEXA-MSP and AD-MSP protein fusions. Association *via* the dimerization domain is shown as a vertical interaction, and association *via* the filament assembly domain is shown as a horizontal interaction. DIMER⁺, indicates dimer-proficient mutants; DIMER⁻, indicates dimer-deficient mutants.

If MSP–MSP association within the two-hybrid system occurs *via* two distinct sites of interaction, then why should a mutation that affects only one site (e.g. the dimerization domain) prevent interaction through the second site (the filament assembly domain)? We believe that wild-type MSP preferentially associates with other wild-type MSP molecules, and this preferential association prevents interaction with mutant MSP. Because heteromeric interactions between LEXA-MSP and AD-MSP molecules are necessary to activate transcription of the *lacZ* reporter gene, preferential self-association among molecules of one fusion protein would effectively prevent interaction with the other fusion protein and manifest itself as an interaction defect. Thus, dimerization-deficient mutants (Figure 7C and G) could not form heterodimers with wild-type MSP, due to mutation, and homotetramer formation of wild-type MSP would preclude interaction with the mutant *via* the filament assembly domain. Dimerization-proficient mutants (Figure 7B and D) would fail to form heterodimers with wild-type MSP (see following), and would also fail to form tetramers, due to the mutation in the filament assembly domain.

Our model requires that mutations in residues outside the dimerization domain must nonetheless prevent heterodimer formation with wild-type MSP. A naturally occurring example of this phenomenon is found in *A. suum*, which contains two isoforms of MSP. The two isoforms, called α and β , differ by only four amino acids at (numbering relative to *C. elegans* MSP; see Figure 6) positions 15 (α , Ser; β , Gly), 16 (Gln; Ser), 55 (Ser; Gly), and 68 (Lys; Ser). None of these residues lies within the dimerization domain, yet only α – α or β – β homodimers are formed in solution. Similarly, missense mutations within the *f* and *g* strands pre-

vent the formation of heterodimers with wild-type MSP, but still allow the formation of mutant homodimers. This observation also indicates a limitation of the two-hybrid system for identifying MSP mutants: only those mutations that prevent heterodimer formation are detectable in our screen.

If all of the mutations recovered are able to disrupt heterodimer formation with wild-type MSP, then why invoke a second site of interaction (i.e. the proposed filament assembly domain)? The data from suppressor analysis (Figure 5) indicate that the dimerization interface cannot be the sole site of interaction. Dimer-proficient mutants could form heterodimers with other dimer-proficient mutants *via* the dimerization interface (Figure 7E); however, dimer-deficient mutants are unable to interact *via* this domain. These results suggest that interaction between dimer-deficient mutants occurs through a second site. The most compelling example is the N126K mutation, which lies within the dimerization interface. We have shown that this mutation prevents homodimer formation *in vitro* (Figure 4); however, LEXA-MSP*(N126K) interacts with AD-MSP*(N126K) *in vivo* (Figure 5). Because the dimer interaction has been disrupted, this association must occur *via* a different site of interaction. MSP is known to polymerize into filaments, so the second site of interaction is most likely a filament assembly domain. Thus, dimer-deficient mutants would interact with either dimer-proficient (Figure 7F and H) or dimer-deficient mutants (Figure 7I) *via* the filament assembly interface. Because truncation mutants that lack the *g* strand but retain the *f* strand associate with the dimer-deficient mutants, we suggest that this interaction is also mediated by the filament assembly domain. This interpretation implies that the *f* strand is the site of a filament assembly interaction.

Ultimately, the mutants isolated in this screen should prove useful in further structural and biochemical analyses of MSP filament assembly. While our data confirm the crystallographic identification of the dimer interface, the putative filament formation domain could represent one of two potential interactions: either the site required for the assembly of MSP into subfilaments, or the one involved in the assembly of helical subfilaments into filaments. Determination of the crystal structure of mutant MSPs and comparison with wild-type might distinguish between the two interactions, as well as identify the precise structural changes caused by the mutations. Also, the use of dimerization-defective MSPs for *in vitro* filament assembly assays could demonstrate conclusively whether or not dimerization of MSP is a necessary prerequisite for filament formation. Finally, these mutants may prove useful in the analysis of other proteins that interact with MSP to regulate filament formation *in vivo*.

Materials and Methods

Plasmids

Two-hybrid plasmids

The *lacZ* reporter plasmids pJK103 and pSH18-34, and the LEXA-BICOID plasmid pRFHM1, have been described (Golemis *et al.*, 1997). The *msp-142* gene was cloned from *C. elegans* genomic DNA by PCR amplification with primers oMSP-56-5' (5'-GGAATTCAC-CATGGCCCAATCCG-3') and oMSP-142-3' (5'-AAA-ACTCGAGTACAAGGAACCTATGGG-3'); primers were designed to generate an *NcoI* site at the start codon, an *EcoRI* site at the 5' end, and a *XhoI* site at the 3' end. The 400 base-pair *msp-142* fragment was amplified by 30 cycles of 30 seconds at 94°C, 30 seconds at 46°C, and 45 seconds at 72°C. The PCR fragment was digested with *EcoRI* and *XhoI* and ligated into plasmid pJG4-5 (Gyuris *et al.*, 1993) digested with *EcoRI* and *XhoI* to create plasmid pAD-MSP. The sequence of the MSP fragment was determined to be identical to the *msp-142* gene sequence. Plasmid pLEXA-MSP was constructed by digestion of pAD-MSP with *EcoRI* and *XhoI*, gel-purification of the 400 base-pair MSP fragment, and ligation into *EcoRI-XhoI*-digested plasmid pGILDA; this plasmid contains the *GAL1* promoter upstream of the *lexA* DNA binding domain (D. Shaywitz, personal communication).

Mutant pAD-MSP* plasmid libraries

Eight independent libraries of pAD-MSP mutants (designated pAD-MSP*) were generated *via* mutagenic PCR (Leung *et al.*, 1989). Each library was constructed from PCR fragments produced under differing mutagenic conditions. The *msp-142* gene was amplified from plasmid pAD-MSP with primers oJG4-5-5' (5'-CTGAGTGGAGATGCCTCC-3') and oJG4-5-3' (5'-CCTTGATTGGAGACTTGAC-3'). For libraries 1 to 4, PCR was performed with one nucleotide (dATP, dCTP, dGTP, or dTTP, respectively) at reduced concentration (0.5 mM) relative to the other dNTPs (2.5 mM). For libraries 5 to 8, nucleotide concentrations were uniform but reactions contained increasing concentrations of MnCl₂ (0, 50, 100, and 150 μM, respectively). All reac-

tions contained 1 ng pAD-MSP template DNA, 0.4 μM each primer, 10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1.5 mM MgCl₂, and 0.001% (w/v) gelatin in a 50 μl volume. PCR reactions consisted of 25 cycles of 30 seconds at 94°C, 30 seconds at 54°C, and 45 seconds at 72°C. Each PCR fragment was purified *via* QIAquick (Qiagen, Chatsworth, CA), digested with *EcoRI* and *XhoI*, gel-purified, and ligated into *EcoRI-XhoI*-digested pJG4-5. Each library produced between 5.4×10^3 and 1.4×10^4 independent clones, with >99.7% of the clones containing inserts.

Mutant pLEXA-MSP* plasmids

Each mutant pLEXA-MSP plasmid (designated pLEXA-MSP*) was constructed from the appropriate pAD-MSP* plasmid by digestion with *EcoRI* and *XhoI*, gel purification of the 400 base-pair MSP fragment, and ligation into *EcoRI-XhoI*-digested plasmid pGILDA.

pET-MSP plasmids

To permit bacterial synthesis of MSP protein, the *msp-142* gene was cloned into the bacterial expression vector pET15b (Novagen, Madison, WI); this construct places MSP under control of the inducible bacteriophage T7 *lac* promoter. Plasmid pET-MSP was constructed by digestion of pAD-MSP with *NcoI* and *XhoI*, gel-purification of the 400 base-pair fragment, and ligation into *NcoI-XhoI*-digested pET15b. Mutant pET-MSP plasmids were constructed in the same manner, using the mutant MSP fragment from the respective mutant pAD-MSP* plasmid.

Two-hybrid screen for pAD-MSP* interaction-defective mutants

Saccharomyces cerevisiae strain EGY48 (genotype *MATα trp1 his3 ura3 6-OP_{lexA}-LEU2*; Gyuris *et al.*, 1993) was first transformed with pLEXA-MSP and the *lacZ* reporter plasmid pSH18-34 or pJK103 to generate Tx-1 or Tx-2, respectively. Transformants were selected on SC-His-Ura plates. Each mutant pAD-MSP* library then was transformed independently into Tx-1 and selected on SC-His-Ura-Trp plates. These transformants were replica-plated onto 5Gal-His-Ura-Trp/X-Gal plates. Potential pAD-MSP* mutants were identified as white colonies after three days of growth. Plasmids were recovered by growing mutant transformants overnight in SC-Trp medium (to maintain selection on pAD-MSP*), isolating DNA from yeast, and transforming *Escherichia coli* strain DH5-α *via* electroporation. Recovery of pAD-MSP* was confirmed by PCR screening with primers oJG4-5-5' and oJG4-5-3'; the presence of a 400 base-pair fragment is diagnostic of the desired plasmid. Potential mutants were retransformed into Tx-2 and rescreened on 5Gal-His-Ura-Trp/X-Gal plates. Mutations were identified by dideoxy-sequencing (Sequenase, United States Biochemical) of one strand and confirmed by automated cycle sequencing (ABI) of the complementary strand.

Two-hybrid screen for suppressor mutants

Yeast strain EGY48 was first transformed with a mutant pLEXA-MSP* plasmid plus the *lacZ* reporter plasmid pSH18-34 or pJK103. Transformants were selected on SC-His-Ura plates. A single colony from each transformation was then transformed with pooled DNA

from mutant pAD-MSP* libraries 1 to 8 and selected on SC-His-Ura-Trp plates. These transformants were replica-plated onto SGal-His-Ura-Trp/X-Gal plates. Potential pAD-MSP* suppressor mutants were identified as blue colonies after three days of growth. Plasmids were recovered and sequenced as described above. pLEXA-MSP* mutants and pAD-MSP* suppressor mutants were retested in all pairwise combinations to determine the specificity of suppression.

β -Galactosidase assays

Yeast transformants containing the indicated plasmids were grown overnight in SC-His-Ura-Trp medium at 30°C. Cells were harvested by centrifugation, resuspended in SGal-His-Ura-Trp at an $A_{600} = 0.1$, grown for eight hours at 30°C, and A_{600} values were recorded. Aliquots of 1 ml were harvested by centrifugation and resuspended in 150 μ l of Z buffer (Miller, 1972), 50 μ l CHCl_3 , and 20 μ l 0.1% (w/v) SDS. Cells were permeabilized by vortex mixing for 30 seconds. β -Galactosidase activity was assayed by the addition of 700 μ l Z buffer containing 1 mg/ml *o*-nitrophenyl- β -D-galactopyranoside (ONPG). Reactions were halted by the addition of 500 μ l 1 M Na_2CO_3 and the reaction time was recorded. The absorbance at 420 nm (A_{420}) was measured and units of activity were calculated using the formula: $(A_{420} \times 1000)/(\text{minutes reaction time} \times A_{600})$. Values represent the mean of three or four independent transformants. Variation between samples was less than 30% of the mean.

Bacterial expression of MSP

pET-MSP wild-type and mutant plasmids were transformed into *E. coli* strain BL21(DE3) (Studier & Moffatt, 1986), which contains the phage T7 RNA polymerase gene under control of the inducible *lacUV5* promoter. Addition of isopropyl- β -D-thiogalactopyranoside (IPTG) stimulates expression of T7 RNA polymerase, which then drives transcription of MSP from the T7 *lac* promoter. pET-MSP plasmids were deleterious to the growth of *E. coli* even in the absence of induction, presumably due to endogenous levels of T7 RNA polymerase (and therefore MSP) expression. Since phage T7 lysozyme has been shown to inhibit the activity of T7 RNA polymerase, the strain was co-transformed with the lysozyme-expressing plasmid pLysS to further decrease basal expression of MSP (Studier, 1991). Transformants were selected on LB medium containing 100 μ g/ml ampicillin and 34 μ g/ml chloramphenicol. Transformants were grown overnight at 37°C in LB medium containing 1 M sorbitol, 2.5 mM betaine (Sigma Chemical, St. Louis, MO), 100 μ g/ml ampicillin and 34 μ g/ml chloramphenicol. Then 100 ml cultures of the same medium were inoculated at $A_{600} = 0.2$ and grown at 25°C until $A_{600} = 0.5$ to 0.6. MSP expression was induced by addition of IPTG to 1 mM concentration. Cells were grown for six hours and harvested by centrifugation. Cell pellets were resuspended in 5 ml extraction buffer (50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol). Cells were lysed by addition of Triton X-100 to 0.1% (v/v) and incubation for five minutes on ice. Lysates were treated with DNase I (1 μ g/ml final concentration) in the presence of MgCl_2 and MnCl_2 (8.0 mM and 2.0 mM, respectively) for 20 minutes at ambient temperature. Lysates were cleared by high speed centrifugation (30 minutes at 45,000 g). The protein extracts were then

dialyzed extensively against column buffer (50 mM NaPO_4 , 150 mM NaCl, 3 mM NaN_3 , pH 7.0) at 4°C. Bradford (1976) assays indicated that total protein concentrations were 10 to 13 mg/ml.

Analytical gel filtration

Monomer/dimer determinations were made essentially as described (Haaf *et al.*, 1996). Gel filtration was performed at ambient temperature using a 16 mm \times 600 mm Sephacryl S-100 High Resolution column (Pharmacia Biotech, Uppsala, Sweden). Glycerol was added to protein extracts at 5% (w/v) final concentration, and 1 ml samples were loaded and eluted at 0.5 ml/minute into 2.5 ml fractions. Proteins in each fraction were concentrated by deoxycholate/trichloroacetic acid precipitation (Mahuran *et al.*, 1983). MSP protein was detected by SDS/PAGE followed by Coomassie blue staining. The column was calibrated using bovine serum albumin ($M = 66 \times 10^3$), carbonic anhydrase ($M = 29 \times 10^3$), and cytochrome *c* ($M = 12.4 \times 10^3$). Void volume was determined using blue dextran ($M \sim 2000 \times 10^3$).

MSP protein sequences

The following MSP protein sequences were used for comparison to identify conserved regions: *C. elegans*, MSP-142 (accession number P53017), MSP-33 (P53019), MSP-3 (P53023), MSP-10 (P05634), MSP-38 (P53020), MSP-63 (PID number g868177), MSP-152 (g868205), MSP-49 (g1166625), MSP-55 (g1208867), and MSP-77 (g1627886) (Burke & Ward, 1983; S. Ward, unpublished data); *A. suum*, MSP- α (P27439) and MSP- β (P27440) (Bennett & Ward, 1986; King *et al.*, 1992); *D. viviparus* MSP (S64873) (Schneider, 1993); *O. vulvulus*, MSP1 (P13262) and MSP2 (P13263) (Scott *et al.*, 1989b); and *G. rostociensis*, MSP1 (P53021), MSP2 (P53022), and MSP3 (P53023) (Novitski *et al.*, 1993). Note that the *A. suum* sequences were edited from D to G at position 113 to correct a typographical error in submission (Tom Roberts, personal communication), and the *D. viviparus* sequence was edited to correct the reading frame (Setterquist & Fox, 1995).

Media and miscellany

Bacterial and yeast media were prepared as described (Ausubel *et al.*, 1997). SC-Trp medium lacks tryptophan, SC-His-Ura lacks histidine and uracil, and SC-His-Ura-Trp lacks all three. Yeast transformations were performed using the high-efficiency LiAc method (Geitz & Schiestl, 1995). Plasmids were recovered from yeast *via* the DNA quick prep method (Hoffman & Winston, 1987).

Acknowledgements

Special thanks to E. Golemis for providing plasmids, yeast strains, and many helpful suggestions for the two-hybrid experiments. We thank D. Shaywitz for providing plasmid pGILDA, and M. Stewart for communicating the crystal coordinates of *A. suum* MSP- α prior to publication. We thank B. Patterson, E. Little and members of the Ward lab for fruitful discussions and critical comments on the manuscript. Special thanks to Bob Speer,

whose senior honors thesis was the genesis of this project. This work was supported by NIH grant GM25243 (S.W.) and an American Cancer Society postdoctoral fellowship PF4199 (H.E.S.).

References

- Amberg, D. C., Basart, E. & Botstein, D. (1995). Defining protein interactions with yeast actin *in vivo*. *Nature Struct. Biol.* **2**, 28–35.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1997). *Current Protocols in Molecular Biology*, Janssen, K., ed., John Wiley & Sons Inc, New York.
- Bennett, K. L. & Ward, S. (1986). Neither a germ-line nor several somatically expressed genes are lost or rearranged during embryonic chromatin diminution in the nematode *Ascaris lumbricoides* var. *suum*. *Dev. Biol.* **118**, 141–147.
- Blackwell, J. R. & Horgan, R. (1991). A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Letters*, **295**, 10–12.
- Bork, P., Holm, L. & Sander, C. (1994). The immunoglobulin fold: Structural classification, sequence patterns and common core. *J. Mol. Biol.* **242**, 309–320.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
- Bullock, T. L., Roberts, T. M. & Stewart, M. (1996). 2.5 Å resolution crystal structure of the motile major sperm protein (MSP) of *Ascaris suum*. *J. Mol. Biol.* **263**, 284–296.
- Burke, D. J. & Ward, S. (1983). Identification of a large multigene family encoding the major sperm protein of *Caenorhabditis elegans*. *J. Mol. Biol.* **171**, 1–29.
- Geitz, R. D. & Schiestl, R. H. (1995). Transforming yeast with DNA. *Methods Mol. Cell. Biol.* **5**, 255–269.
- Golemis, E. A., Serebriiskii, I., Gyuris, J. & Brent, R. (1997). Interaction trap/two-hybrid system to identify interacting proteins. In *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds.), John Wiley & Sons Inc, New York.
- Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell*, **75**, 791–803.
- Haaf, A., Butler, P. J. G., Kent, H. M., Fearnley, I. M., Roberts, T. M., Neuhaus, D. & Stewart, M. (1996). The motile major sperm protein (MSP) from *Ascaris suum* is a symmetric dimer in solution. *J. Mol. Biol.* **260**, 251–260.
- Hoffman, C. F. & Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, **57**, 267–272.
- Holmgren, A. & Branden, C.-I. (1989). Crystal structure of chaperonin protein PapD reveals an immunoglobulin fold. *Nature*, **342**, 248–251.
- Holmgren, A., Kuehn, M. J., Branden, C.-I. & Hultgren, S. J. (1992). Conserved immunoglobulin-like features in a family of periplasmic pilus chaperones in bacteria. *EMBO J.* **11**, 1617–1622.
- Italiano, J. E., Jr, Roberts, T. M., Stewart, M. & Fontana, C. A. (1996). Reconstitution *in vitro* of the motile apparatus from the amoeboid sperm of *Ascaris* shows that filament assembly and bundling move membranes. *Cell*, **84**, 105–114.
- King, K. L., Stewart, M., Roberts, T. M. & Seavy, M. (1992). Structure and macromolecular assembly of two isoforms of the major sperm protein (MSP) from the amoeboid sperm of the nematode, *Ascaris suum*. *J. Cell Sci.* **101**, 847–857.
- King, K. L., Essig, J., Roberts, T. M. & Moerland, T. S. (1994). Regulation of the *Ascaris* major sperm protein (MSP) cytoskeleton by intracellular pH. *Cell Motil. Cytoskel.* **27**, 193–205.
- L'Hernault, S. W. (1997). Spermatogenesis. In *C. elegans II* (Riddle, D. L., Blumenthal, T., Meyer, B. J. & Priess, J. R., eds.), pp. 271–294, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Leung, D. W., Chen, E. & Goeddel, D. V. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique*, **1**, 11–15.
- Lorenz, M., Popp, D. & Holmes, K. C. (1993). Refinement of the F-actin model against X-ray fiber diffraction data by the use of a directed mutation algorithm. *J. Mol. Biol.* **234**, 826–836.
- Ma, J. & Ptashne, M. (1987). A new class of yeast transcriptional activators. *Cell*, **51**, 113–119.
- Mahuran, D., Clements, P., Carrella, M. & Strasberg, P. M. (1983). A high recovery method for concentrating microgram quantities of protein from large volumes of solution. *Anal. Biochem.* **129**, 513–516.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nelson, G. A., Roberts, T. M. & Ward, S. (1982). *Caenorhabditis elegans* spermatozoan locomotion: amoeboid movement with almost no actin. *J. Cell Biol.* **92**, 121–131.
- Novitski, C. E., Brown, S., Chen, R., Corner, A. S., Atkinson, H. J. & McPherson, M. J. (1993). Major sperm protein genes from *Globodera rostochiensis*. *J. Nematol.* **25**, 548–554.
- Roberts, T. M. & King, K. L. (1991). Centripetal flow and directed reassembly of the major sperm protein (MSP) cytoskeleton in the amoeboid sperm of the nematode, *Ascaris suum*. *Cell Motil. Cytoskel.* **20**, 228–241.
- Roberts, T. M., Sepsenwol, S. & Ris, H. (1989). Sperm motility in nematodes: crawling movement without actin. In *The Cell Biology of Fertilization*, pp. 41–60, Academic Press, Inc., New York.
- Sayle, R. (1996). RasMol, v. 2.6 (freeware).
- Schnieder, T. (1993). The diagnostic antigen encoded by gene fragment DV3-14: a major sperm protein of *Dictyocaulus viviparus*. *Int. J. Parasitol.* **23**, 383–389.
- Scott, A. L., Dinman, J., Sussman, D. J. & Ward, S. (1989a). Major sperm protein and actin genes in free-living and parasitic nematodes. *Parasitology*, **98**, 471–478.
- Scott, A. L., Dinman, J., Sussman, D. J., Yenbutr, P. & Ward, S. (1989b). Major sperm protein genes from *Onchocerca volvulus*. *Mol. Biochem. Parasitol.* **36**, 119–126.
- Sepsenwol, S., Ris, H. & Roberts, T. M. (1989). A unique cytoskeleton associated with crawling in the amoeboid sperm of the nematode *Ascaris suum*. *J. Cell Biol.* **108**, 55–66.
- Setterquist, R. A. & Fox, G. E. (1995). *Dictyocaulus viviparus*: nucleotide sequence of Dv3-14. *J. Parasitol.* **25**, 137–138.

- Stewart, M., King, K. L. & Roberts, T. M. (1993). Crystallization of the motile major sperm protein (MSP) of the nematode *Ascaris suum*. *J. Mol. Biol.* **232**, 298–300.
- Stewart, M., King, K. L. & Roberts, T. M. (1994). The motile major sperm protein (MSP) of *Ascaris suum* forms filaments constructed from two helical subfilaments. *J. Mol. Biol.* **243**, 60–71.
- Studier, F. W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* **219**, 37–44.
- Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130.
- Theriot, J. A. (1996). Worm sperm and advances in cell locomotion. *Cell*, **84**, 1–4.
- Varkey, J. P., Jansma, P. L., Minniti, A. N. & Ward, S. (1993). The *Caenorhabditis elegans spe-6* gene is required for major sperm protein assembly and shows second site non-complementation with an unlinked deficiency. *Genetics*, **133**, 79–86.
- Ward, S. & Klass, M. (1982). The location of the major protein in *Caenorhabditis elegans* sperm and spermatocytes. *Dev. Biol.* **92**, 203–208.
- Ward, S., Argon, Y. & Nelson, G. A. (1981). Sperm morphogenesis in wild-type and fertilization-defective mutants of *Caenorhabditis elegans*. *J. Cell Biol.* **91**, 26–44.

Edited by A. R. Fersht

(Received 17 November 1997; received in revised form 10 March 1998; accepted 20 March 1998)