Cloning, Zn^{2+} Binding, and Structural Characterization of the Guanine Nucleotide Exchange Factor Human Mss4†

Hongtao Yu and Stuart L. Schreiber*

Howard Hughes Medical Institute, Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received February 27, 1995; Revised Manuscript Received April 24, 1995⊗

ABSTRACT: The Sec4/Ypt1/Rab family of small GTP-binding proteins are involved in the regulation of intracellular vesicular transport. A rat gene, *mss4*, that encodes a guanine nucleotide exchange factor (GEF) for Sec4 was recently cloned by its ability to rescue defects in protein transport of a yeast temperature-sensitive (ts) mutant, *sec4−8*. We describe herein the cloning, bacterial expression, and biochemical characterization of human Mss4. As expected, both the cDNA and its encoded amino acid sequences are highly conserved between the human and rat *mss4*. Soluble and functional Mss4 protein was obtained by expressing the gene as a glutathione-*S*-transferase fusion protein in *Escherichia coli*. Subsequent biochemical analysis revealed that Mss4 binds 1 equiv of Zn^{2+}, and zinc is essential for the stability of the protein. Utilizing multidimensional heteronuclear NMR techniques, we assigned most of the ^{1}H, ^{15}N, and ^{13}C resonances of this 14-kDa protein. Its secondary structure was also deduced from slowly exchanging amide protons, characteristic NOEs, and $^{3}J_{NH\text{-}C_{\alpha}H}$ coupling constants. The protein contains a central seven-stranded antiparallel β sheet, flanked by two small β sheets. Many resonances pertaining to a loop region of the molecule cannot be identified, suggesting that it might be involved in local movements. These studies provide the first structural insights into a protein possessing GEF activity.

Members of the Ras superfamily of small GTPases have been implicated in the control of many cellular processes, including cell proliferation, protein synthesis, and vesicular traffic (Hall, 1990; Bourne et al., 1991; Pfeffer, 1994). The roles of Ras proteins in vesicle transport were first discovered when one of the yeast secretion (sec) mutants, Sec4, was isolated (Salminen & Novick, 1987). Mutations within the *SEC4* gene block vesicle transport from the Golgi to the plasma membrane and cause accumulation of secretory vesicles in a temperature-sensitive manner. Cloning and sequencing of the *SEC4* gene revealed that the Sec4 protein shares about 30% amino acid sequence identity with Ras. Another yeast protein, Ypt1, which shares 45% sequence identity with Sec4, was shown to regulate transport between the endoplasmic reticulum (ER) and the Golgi (Segev et al., 1988). Several mammalian proteins, termed Rabs, were subsequently isolated and found to be closely related to Sec4 and Ypt1 (Novick & Brennwald, 1993; Fischer von Mollard et al., 1994). As proteins belonging to the Sec4/Ypt1/Rab subfamily of GTPases act on virtually every stage of secretion and endocytosis, they were initially thought to target specific transporting vesicles to their fusion membranes. However, chimeric proteins constructed from Sec4 and Ypt1 were not sufficient to cause missorting of these vesicles (Brennwald & Novick, 1993; Dunn et al., 1993). More recently, it has been postulated that a family of SNAREs determine the pairwise matching of donor and acceptor membranes (Rothman, 1994). One possible effector of Sec4 was shown to be Sec9, which is a component of a yeast SNARE complex (Brennwald et al., 1994). In addition, a Rab protein, Sly1, is required for the assembly of SNARE complexes in yeast (Søgaard et al., 1994). Therefore, Rabs seem to be involved in the docking and fusion events of transport vesicles, rather than their targeting (Ferro-Novick & Jahn, 1994).

As a result of prenylations near their C-termini, Rabs are predominantly associated with membranes of transporting vesicles or their targeting compartments (Goud et al., 1988; Pfeffer, 1994). The recycling of functional forms of Rab proteins is facilitated by a set of auxiliary proteins, including GAPs,[1] GEFs, and GDIs (Novick & Brennwald, 1993). Moya et al. have recently identified a protein called DSS4 that is a GEF for Sec4 in a suppressor screening of *sec4−8* (Moya et al., 1993). Meanwhile, a mammalian homolog of DSS4 (27% identity with DSS4), called Mss4, was cloned in a similar fashion (Burton et al., 1993). Both proteins are able to stimulate GTP−GDP exchange in Sec4, Ypt1, and Rab3a, but neither could act on the yeast Ras2 protein. As Mss4 and DSS4 share no significant sequence homology with GEFs for Ras, Rho, or Ran, they represent a new family of guanine exchange prtoteins (Boguski & McCormick, 1993; Feig, 1994). More recently, Mss4 was shown to bind selectively to a subset of genetically related Rab proteins

† Supported by a grant from the National Institute of Health (GM 44993) to S.L.S. S.L.S. is an Investigator at the Howard Hughes Medical Institute.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1995.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two-dimensional; 3D, three-dimensional; HMQC, heteronuclear multiple-quantum correlation; HSQC, heteronuclear single-quantum correlation; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; ppm, parts per million; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation; Mss4, mammalian suppressor of Sec4; DSS4, dominant suppressor of Sec4; GEF, guanine nucleotide exchange factor; GAP, GTPase-activating protein; GDI, GDP dissociation inhibitor; SNARE, SNAP receptor; SNAP, soluble NSF-associated protein; PCR, polymerase chain reaction; PMPS, *p*-(hydroxymercuri)-phenylsulfonate; PAR, 4-(2-pyridylazo)resorcinol; IPTG, isopropyl β-D-thiogalactopyranoside.

0006-2960/95/0434-9103$09.00/0

and enhance their rate of guanine nucleotide exchange (Burton et al., 1994). Furthermore, Mss4 co-immunoprecipitated with Rab3a in a brain extract by anti-Rab3a antibodies (Burton et al., 1994). Consistent with the observation that Rab3a is important for the recruitment of synaptic vesicles for exocytosis (Geppert et al., 1994), Mss4 facilitated neurotransmitter release when injected into the squid giant nerve terminal (Burton et al., 1994). Thus, it is becoming increasingly evident that Mss4 specifically regulates the functions of certain Rab GTPases *in vivo* by promoting their exchange of GTP and GDP.

We have now cloned human Mss4 (hMss4) by screening a human brain cDNA library with radioactive oligonucleotide probes. The amino acid sequences of human and rat proteins are highly similar except their N-terminal regions. Recombinant hMss4 was obtained by overexpression of the gene as a GST fusion protein in *Escherichia coli*, cleavage of the fusion protein with thrombin, and subsequent separation of the mixture with gel filtration chromatography. The bacterially expressed protein is fully active in its ability to stimulate the exchange of GTP for GDP in Sec4 and Ypt1. Furthermore, we have shown that hMss4 can bind 1 mol equiv of Zn^{2+} as determined by a colorimetric assay. The zinc ion is a critical structural element of the protein. Since structural information concerning guanine nucleotide exchange factors does not currently exist, and since hMss4 is a protein with low molecular weight (14 kDa), we set out to determine its tertiary structure by multidimensional NMR. Toward this goal, we have now completed the assignment of ^1H, ^{15}N, and ^{13}C resonances of the protein. In addition, the secondary structure of hMss4 has also been determined on the basis of hydrogen bonds, crucial NOEs, and $^3J_{NH\text{-}C_\alpha H}$ coupling constants. hMss4 mainly consists of β structure, including a major seven-stranded antiparallel sheet, a small two-stranded sheet, and a small three-stranded sheet. There are also two short helical segments located at residues 74−77 and 117−120, respectively. Residues 49−58 appear to be highly dynamic, as evidenced by the absence of many of their amide NH and ^{15}N correlation peaks in an HSQC spectrum.

MATERIAL AND METHODS

Cloning of Human Mss4. The rat Mss4 cDNA was cloned by screening a λgt11 rat brain cDNA library with a ^{32}P-labeled probe corresponding to codons 60−74 of the published sequence (Burton et al., 1993). The open reading frame (ORF) of the rat Mss4 gene was amplified by PCR. The PCR product was subsequently labeled with ^{32}P using the Klenow fragment of DNA polymerase with random hexamers as primers (Boehringer Mannheim). The labeled DNA mixtures were used to probe a λZAPII human brain cDNA library (Strategene). Two positives were detected from screening five plates with 5×10^4 plaques/plate, and the positive plaques were purified. Excision of the phagemids by the R408 helper phages resulted in pBS vectors that were propagated and sequenced using the Sequenase kit (USB).

Bacterial Expression and Purification of Human Mss4. The human *mss4* gene was amplified by PCR. After digestion with *Bam*HI and *Eco*RI, the resulting PCR products were subcloned into the expression vector pGEX-2T, which was tranferred into *E. coli* strain BL21. The bacteria in LB media were grown to midlog phase at 37 °C and induced with 1 mM IPTG for 8 h at 30 °C. Harvested cells were lysed by sonication and pelleted by centrifugation. The supernatant was loaded onto a glutathione−agarose column (Sigma). After extensive washes with Tris-HCl buffer, the fusion protein was eluted with Tris-HCl buffer containing 5 mM reduced glutathione and cleaved using thrombin (Sigma). Homogeneous Mss4 protein was obtained after passing the cleavage mixture through a Sephacryl S-100 gel filtration column (Pharmacia). The ^{15}N- and ^{13}C-labeled Mss4 was produced by growing cells in M9 minimal medium with $^{15}NH_4Cl$ and [^{13}C]glucose as the sole nitrogen and carbon sources, respectively. Cells grown in the minimal medium were induced by adding IPTG and $ZnCl_2$ to final concentrations of 1 and 0.1 mM, respectively.

Guanine Nucleotide Exchange Assay. The yeast *SEC4* gene was obtained from yeast genomic DNA using PCR with the primers 5′-AGC GGA TCC ATG TCA GGC TTG AGA ACT GTT TCT-3′ and 5′-GTG GAA TTC TCA ACA GCA ATT TGA TTT AGA ACT GTT-3′. The PCR fragment was subcloned into the pGEX-2T vector. The Sec4 protein was produced and purified as described above. The Ypt1 protein was a generous gift from Dr. Morgen Søgaard in the Rothman Laboratory at the Memorial Sloan Kettering Cancer Center. Ypt1p or Sec4p was incubated at 30 °C for 30 min with [^3H]GDP in a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 2 mM EDTA, 1 mM DTT, 50 mM NaCl, and 50 μg/mL of BSA. The hMss4 protein and GTP were then added to the reaction mixture. Aliquots were taken at the indicated times and spotted on nitrocellulose filters. After extensive washes with a buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, and 1 mM $MgCl_2$, radioactivity retained on the filters was quantified by a scintillation counter (Beckman). For the control experiment, an equal volume of buffer was added to the reaction instead of the hMss4 protein.

Zinc-Binding Assay. The PAR−PMPS Zn^{2+}-binding assay was adapted from Giedroc et al. (Giedroc et al., 1986). Briefly, recombinant hMss4 protein was exchanged into the TNG buffer (10 mM Tris-HCl, 200 mM NaCl, and 5% glycerol). PAR was added to the protein solution to a final concentration of 100 μM. The protein-bound Zn^{2+} ions were released from the protein with the addition of an excess of PMPS (Sigma). After a plateau of the absorption at 500 nm was reached, the concentration of Zn^{2+}-PAR was calculated with an extinction coefficient of 6.6×10^4 M^{-1} cm^{-1}.

Preparation of NMR Samples. The purified Mss4 was dialyzed against a solution of 50 mM KCl and 1 mM DTT. An Amicon stir cell unit was used to concentrate the protein sample to a concentration of about 3 mM. The concentrated sample was frozen with liquid nitrogen, lyophilized, and redissolved in a buffer containing 40 mM deuterated Tris-HCl (pH 6.5). The final NMR samples contained 3 mM hMss4 protein, 40 mM Tris-HCl, 100 mM KCl, 2 mM DTT, and 0.02% NaN_3. The D_2O samples were prepared by lyophilizing the H_2O sample and dissolving it in 99.999% D_2O.

NMR Spectroscopy. All NMR data were recorded at 303 K on a Bruker DMX500 spectrometer. Quadrature detection in the indirectly detected dimensions was achieved with the TPPI method. The data were processed with the FELIX software with appropriate apodization, base line correction,

and zero-filling to yield real 2D 1K × 1K or 3D 512 × 256 × 256 matrices after reduction (Hare, 1991). For the experiments recorded in H_2O buffer, presaturation during the 0.8-s relaxation delay was used to suppress the intense H_2O signal. Complete removal of the H_2O signal was achieved by deconvoluting the fids before Fourier transform. Unless otherwise mentioned, the spectral widths for ^1H and ^{15}N were 7002 and 2000 Hz, respectively.

The backbone assignment of ^1H, ^{15}N, and ^{13}C was made with the HNCA, HN(CO)CA, and HNCO spectra performed on a $^{15}N/^{13}C$ doubly labeled hMss4 sample in 92% H_2O/8% D_2O (Kay et al., 1990; Bax & Ikura, 1991). The ^{15}N decoupling during detection was achieved by the GARP sequence for these experiments. The HNCA and HN(CO)-CA spectra were acquired with a ^{13}C spectral width of 4000 Hz and with 32* complex points in t_1 (^{15}N), 50* points in t_2 (^{13}C), and 512* points in t_3. In the case of the HNCO experiment, a 32* × 64* × 512* data set was recorded with a ^{13}C spectral width of 2000 Hz. For all three experiments, 16 scans were taken for each increment.

In order to obtain side chain assignments and NOE identification, the HCCH-TOCSY (τ = 25 ms), $^1H/^{15}N$ TOCSY HSQC (τ = 65 ms), $^1H/^{15}N$ NOESY-HMQC (τ = 100 ms), and $^1H/^{13}C$ NOESY-HMQC (τ = 100 ms) spectra were performed (Bax et al., 1990; Marion et al., 1989; Zuiderweg & Fesik, 1989). These spectra were acquired with 8 scans/increment. The GARP pulse sequence was used to decouple ^{15}N or ^{13}C during the detection period. Both the $^1H/^{15}N$ TOCSY-HSQC and $^1H/^{15}N$ NOESY-HMQC spectra contained 90*(^1H) × 80*(^{15}N) × 512* points in t_1, t_2, and t_3, respectively. The $^1H/^{13}C$ NOESY-HMQC and HCCH-TOCSY spectra were obtained with a D_2O sample. The ^1H carrier frequency in the detection period was set in the center of the aliphatic protons, while the residual HDO signal was suppressed by off-resonance Gaussian shaped pulses. The ^{13}C spectral width for the two experiments was 8993 Hz.

The NOEs identified in the 3D NOESY spectra were calibrated by a 50-ms 2D homonuclear NOESY spectra. Whenever overlaps occurred, the NOEs were regarded as weak. The $^3J_{NH\text{-}C_\alpha H}$ coupling constants were measured using an HMQC-J spectrum recorded on the ^{15}N-labeled Mss4 sample in H_2O (Kay & Bax, 1990). Proton−deuterium exchange of amide protons was studied on a sample of ^{15}N-labeled Mss4 lyophilized from H_2O buffer and dissolved in pure D_2O. Three hours following the dissolution of the protein into D_2O, an HSQC spectrum was taken. The amide protons still visible in the processed HSQC spectrum were considered slowly exchanging (Bax et al., 1983).

RESULTS

Cloning, Expression, Purification, and Characterization of Human mss4. As described above, the two positive clones obtained from screening a human cDNA library contain inserts with lengths of about 600 bp and 1 kbp, respectively. The ORFs of the two clones are identical, while the larger insert possesses a much longer 3′-untranslated region. The cDNA sequence and its encoded amino acid sequence of the longer clone are shown in Figure 1a. Because there are no stop codons upstream of the initiation methionine, this clone is incomplete. However, comparison of the human and rat Mss4 protein sequences clearly shows that the human cDNA encodes the full-length Mss4 protein (Figure 1b). The human and rat Mss4 proteins differ by only 11 residues, five of which are homologous substitutions.

The ORF of the human *mss4* gene was amplified by PCR, and the resulting PCR products were subcloned into the expression vector pGEX-2T. A GST-hMss4 fusion protein was produced by overexpressing the pGEX-hMss4 vector in *E. coli* strain BL21. The fusion protein was then purified by glutathione−agarose beads and cleaved using thrombin. Homogeneous hMss4 protein was obtained after passing the cleavage mixture through a Sephacryl S-100 gel filtration column. The purified recombinant hMss4 protein was fully active as demonstrated by the guanine nucleotide exchange assay (Figure 2). hMss4 produced a 5-fold enhancement in the GDP dissociation rates in the Sec4 and Ypt1 proteins.

hMss4 contains two CysXXCys motifs that are 80 residues apart, suggesting that it might be a metal-binding protein. Although the cysteine pairs of conventional zinc finger proteins are normally separated by about 15 amino acids (Berg, 1995), other Zn^{2+}-binding proteins that contain more distal cysteine clusters have also been found. For example, the DNA-binding domain of *E. coli* Ada protein uses four cysteines (Cys 38, Cys 42, Cys 69, and Cys 72) to coordinate a Zn^{2+} ion (Myers et al., 1993). The transcription factor TFIIS also binds Zn^{2+} through a novel fold, designated the zinc ribbon (Qian et al., 1993). The first indication that the hMss4 protein might bind a metal came from a failed attempt to concentrate the protein in a buffer containing phosphate. Under these conditions, the protein was poorly soluble and precipitated slowly. However, it became stable when exchanged into Tris-HCl buffer of the same pH. In addition, we could not obtain soluble hMss4 protein by growing cells in M9 minimal medium that did not contain Zn^{2+}. Supplementing the culture with 0.1 mM of $ZnCl_2$ prior to induction led to a yield of soluble hMss4 protein as high as that of the rich medium. Encouraged by these observations, we carried out a PMPS−PAR Zn^{2+}-binding assay on the hMss4 protein produced from either the rich medium (LB) or the M9 medium with 0.1 mM Zn^{2+} (Giedroc et al., 1986). PMPS can react with the Zn^{2+}-coordinating thiol groups in proteins and cause quantitative release of protein-bound zinc ion. In the case of hMss4, 5 equiv of PMPS were necessary to react with the thiol groups of the five cysteines in the protein. Zinc released by the PMPS reaction can be titrated with the Zn^{2+}-binding dye PAR. Zn(II)PAR_2 has an absorption maximum at 500 nm with an extinction coefficient of 6.6 × 10^4 M^{-1} cm^{-1}. The maximum absorption at 500 nm measured following the titration of an excess of PMPS can therefore be used to calculate the zinc content of the protein. The hMss4 protein purified from the LB medium was found to contain 1.06 equiv of Zn^{2+} ion, while protein from the M9 medium supplemented with 0.1 mM $ZnCl_2$ contained 1.05 equiv of zinc. Furthermore, preliminary 3D structure calculations suggest that Cys 23, Cys 26, Cys 94, and Cys 97 are responsible for coordinating Zn^{2+} (Yu and Schreiber, unpublished results). Thus, the Zn^{2+} ion is essential for maintaining the structural integrity of Mss4. It remains to be seen whether Zn^{2+} is also crucial for its function. Mutations targeted at the Zn^{2+}-coordinating cysteines are expected to cause the unfolding of the protein. Therefore, it will be difficult to differentiate the structural and functional roles of Zn^{2+} by using site-directed mutagenesis alone.

Sequential Assignment. hMss4 is a well-behaved protein in terms of all NMR criteria. It is soluble up to 3 mM

a

```
   1  CCGGCTGCGGCGATGGAACCAGCGGACGAGCCGAGCGAGTTAGTGTCAGCCGAGGGCCGA
   1            M  E  P  A  D  E  P  S  E  L  V  S  A  E  G  R
  61  AACCGGAAGGCGGTGCTGTGCCAGCGTTGCGGCTCCCGGGTGCTGCAGCCAGGGACCGCT
  17  N  R  K  A  V  L  C  Q  R  C  G  S  R  V  L  Q  P  G  T  A
 121  CTCTTCTCTCGCCGACAGCTTTTCCTTCCCTCCATGAGAAAGAAGCCAGCTCTGTCTGAC
  37  L  F  S  R  R  Q  L  F  L  P  S  M  R  K  K  P  A  L  S  D
 181  GGCAGCAATCCTGACGGCGATCTCCTCCAGGAACACTGGCTGGTTGAGGACATGTTCATT
  57  G  S  N  P  D  G  D  L  L  Q  E  H  W  L  V  E  D  M  F  I
 241  TTTGAGAATGTGGGCTTCACCAAGGACGTGGGCAACATCAAGTTTCTGGTCTGCGCAGAC
  77  F  E  N  V  G  F  T  K  D  V  G  N  I  K  F  L  V  C  A  D
 301  TGTGAAATTGGACCAATTGGCTGGCATTGCCTAGATGACAAGAACAGTTTCTATGTGGCC
  97  C  E  I  G  P  I  G  W  H  C  L  D  D  K  N  S  F  Y  V  A
 361  TTGGAACGAGTTTCCCATGAGTAACTGAGGGGAGGGGTACTCAGCTCCATCTCCAAAGAT
 117  L  E  R  V  S  H  E  *
 421  AAACCTACTCCCCACAAGAACTGGCCTTTAATGTGGTATAACTGTTCCGCTGCCTTCTTG
 481  TCTGTGTGCTAATATAAATACTGAGTACCAGCATGTCCATTTGAACATGCAAAGGGTTAA
 541  TCCTGCTTCCTAAAGCCTCAAGTACATGCCTCCTGCTTAGTTCACTTTGTATCACATTTC
 601  CTAAGCTCCCTTTTCCCCCAGTTTTGGGACACTGTGCTTACCTCCAAAAATCTCATCTCT
 661  TCCCTGGCATTCTCCCTAGGCTCTGTTTTGCCCAGGGCTCCCGCTTTTTCTTGCTCTAGA
 721  GGAGCAGTATTCAACCTTTTAGCTATGATGACACATAACAAAAGATGTTTATGTACTAAT
 781  AGTTGAAATCTGCCTTTTTCTCATTCAAGAAGGCATACAAATATCTGAGAGTGACTTTGT
 841  TGTATGGCTACCCTTGTGATCTACAGTAATTTATTCTTTCTAAAAGTAAAGCATTCTCAA
 901  AACAAAAAAAAAAAAAAAAGG
```

b

```
Human Mss4   1 MEPADEPSELVSAEGRNRKAVLCQRCGSRVLQPGTALFSRRQLFLPSMRK 50
               |||.: ..||||||||||||||||||||||||||||||||||||||||||
  Rat Mss4   1 MEPCELQNELVSAEGRNRKAVLCQRCGSRVLQPGTALFSRRQLFLPSMRK 50
                    : .. .| || :. .: ..  :..::. ...: |  .|:     .
Yeast DSS4   1       MSKATCSFEGCHSAVITIN..DDNIINLPEQVHSEFKLL.....E 38

Human Mss4  51 KPALSDGSNPDGDLLQEHWLVEDMFIFENVGFTKDVGN             88
               ||.| |||||||||:|:|||||:|||||||||||||||
  Rat Mss4  51 KPDLVDGSNPDGDVLEEHWLVNDMFIFENVGFTKDVGN             88
               ...: |:. .::::.|    :| |:: |:|||..::::..
Yeast DSS4  39 NRTMRDATPSESNFL....VVPDVWDFDNVGVSREIPSSILGDLSDKSDF 84

Human Mss4  89               IKFLVCADCEIGPIGWHCLDDKNSFYVALERVSHE 123
                             :|||||||||||||||||||||||||||||||||
  Rat Mss4  89               VKFLVCADCEIGPIGWHCLDDKNSFYVALERVSHE 123
                             :|:|:||||: ||||  |  :...   . ||| |
Yeast DSS4  85 VFEYGNSSWKIKKCLKYLICADCDKGPIGIICKVQDQT..KNEERVLHL 131
```

FIGURE 1: (a) Nucleotide and amino acid sequence of human *mss4* cDNA. (b) Sequence alignment of human Mss4, rat Mss4, and DSS4. The identical residues are represented by vertical bars, and conserved substitutions are shown with double dots. The asterisk indicates the Asp to Gly mutation that occurs in *DSS4−1* (Moya et al., 1993).

concentration at pH 6.5 and stable at 30 °C for months. The proton line widths for this protein are in the range of 14−16 Hz. Due to this behavior, we were able to obtain data on hMss4 of excellent quality. An amide ^{1}H- and ^{15}N-based assignment strategy was then utilized to accomplish the sequential assignment. Out of 135 expected ^{1}H/^{15}N correlations of Mss4, 124 are apparent in the HSQC spectrum (Figure 3). These cross peaks were also well-dispersed along both axes. Therefore, each cross peak was considered to represent a unique spin system that was numbered in the order of descending amide ^{1}H chemical shifts. Because hMss4 has relatively sharp ^{1}H line widths, magnetization transfers from NHs up to C_γHs were observed in the ^{1}H/^{15}N TOCSY-HMQC spectrum. This allowed us to correlate the chemical shifts of C_αHs, C_βHs, and in some cases C_γHs and C_δHs to the amide ^{1}H/^{15}N resonance pairs of many designated spin systems. Furthermore, on the basis of the values and patterns of the side chain ^{1}H resonances, we tentatively assigned more than 80% of the spin systems to a certain type of amino acid. Long segments of sequential connections were then made by analyzing NOEs observed in the ^{1}H/^{15}N NOESY-HMQC spectrum. An example is given in Figure 4 that shows the sequential NOEs from Phe 113 to Glu 123. These assignments were confirmed by the HNCA and HN(CO)CA spectra. The HNCA spectrum showed nearly all of the expected intraresidue amide ^{1}H/^{15}N and $^{13}C_\alpha$ correlations. In many cases, the HNCA spectrum also displayed correlations between the amide ^{1}H/^{15}N and the $^{13}C_\alpha$ of the proceeding residue, while the rest of these sequential correlations were provided by the HN(CO)CA experiment.

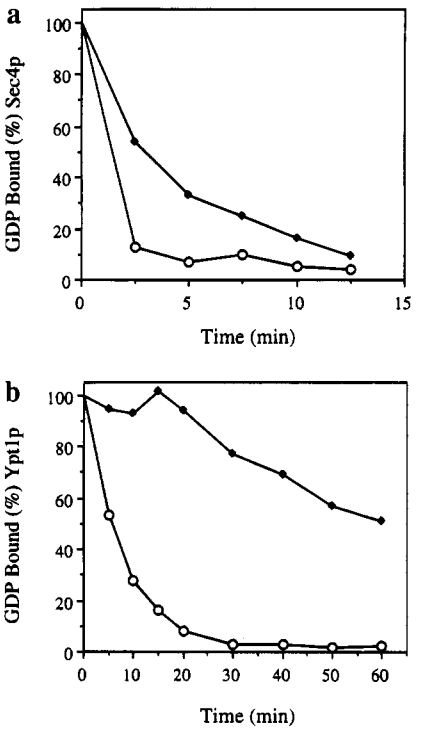

FIGURE 2: GEF assays of recombinant hMss4 with Sec4 (a) and Ypt1 (b) as targets. Open circles represent the hMss4 catalyzed data points; solid diamonds stand for the control data points. Each datum point is the average of two measurements.

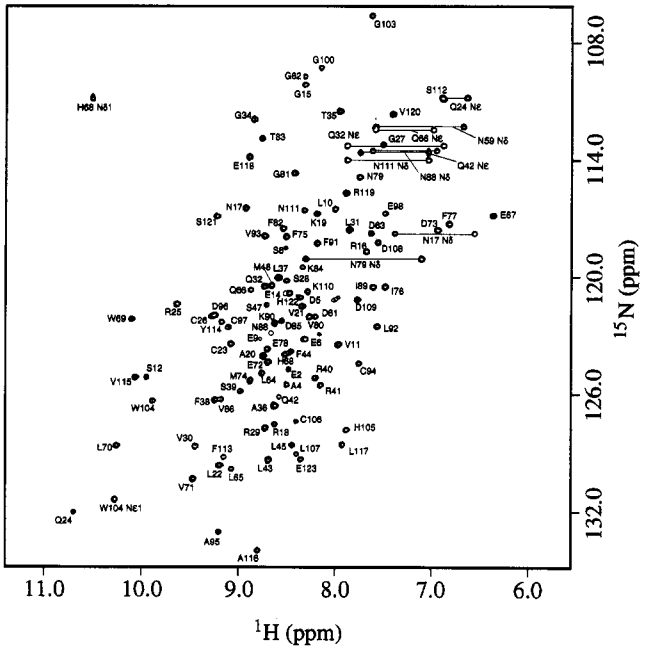

FIGURE 3: HSQC spectrum of hMss4. Peaks for the backbone amide groups are labeled by the residues that give rise to the correlations. Peaks for the side chain amides of Asn and Gln are connected by horizontal lines and labeled with their assignments. The correlation peak for the δ1 proton and ^{15}Nδ1 of His 68 was folded back along the ^{15}N dimension. The actual ^{15}N chemical shift for Nδ1 is 151.4 ppm.

The through-bond connectivities made by a common $^{13}C_\alpha$ chemical shift confirmed the existing assignments and enabled further sequential backbone connections of hMss4, including prolines. Figure 5 displays the HNCA connections for the same region shown in Figure 4. After the sequential assignments of the backbone were established, the HCCH-

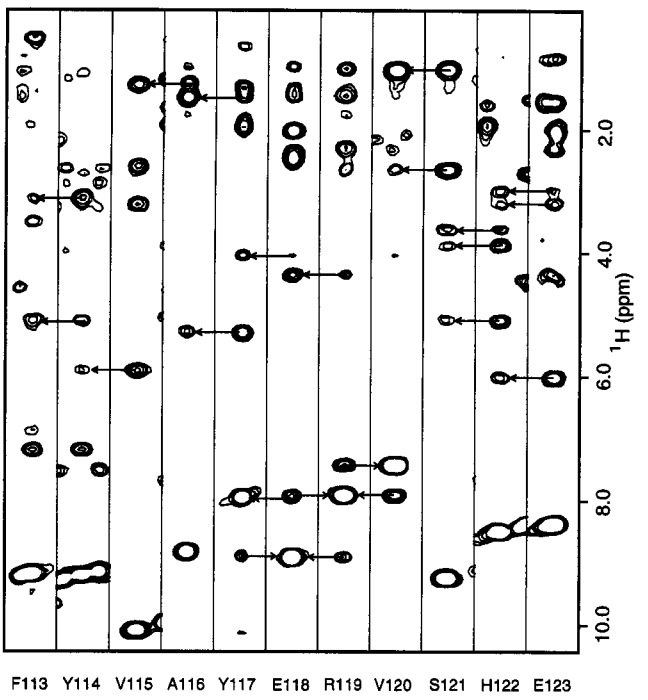

FIGURE 4: Strips from the ^{15}N/^{1}H NOESY-HMQC spectrum, illustrating the sequential through-space connections for residues 113−123. Arrows indicate the direction of the sequential connectivity from the C- to N-terminus.

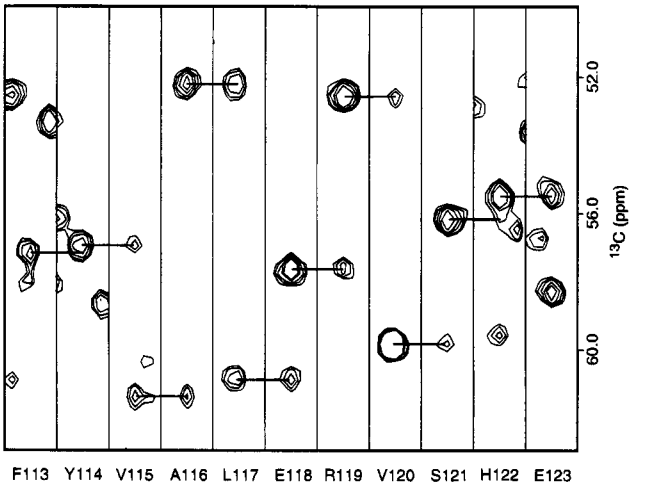

FIGURE 5: Selective slices from the HNCA spectrum. The through-bond connections on the basis of the $13C_\alpha$ chemical shift are shown for residues 113−123. Peaks with weaker intensities are sequential correlations.

TOCSY spectrum was analyzed to identify the missing side chain ^{1}H and ^{13}C resonances. We were not able to assign resonances belonging to residues 49−58, largely due to the absence of their amide ^{1}H/^{15}N cross peaks in the HSQC spectrum. It is possible that this region of the protein is involved in local motions with a medium rate on the NMR time scale. A listing of the ^{1}H, ^{15}N, and ^{13}C chemical shifts of hMss4 is provided as supporting information.

Secondary Structure Analysis. Evidence at the early stage of this study suggested that hMss4 might be comprised mainly of β sheet structures. First, the CD spectrum of hMss4 showed characteristic signals of a β sheet (data not shown). Second, many C_αHs of hMss4 have chemical shifts downfield of the H_2O signal, which is a good indication of β structure. Furthermore, the ^{13}C-dispersed NOESY spec-

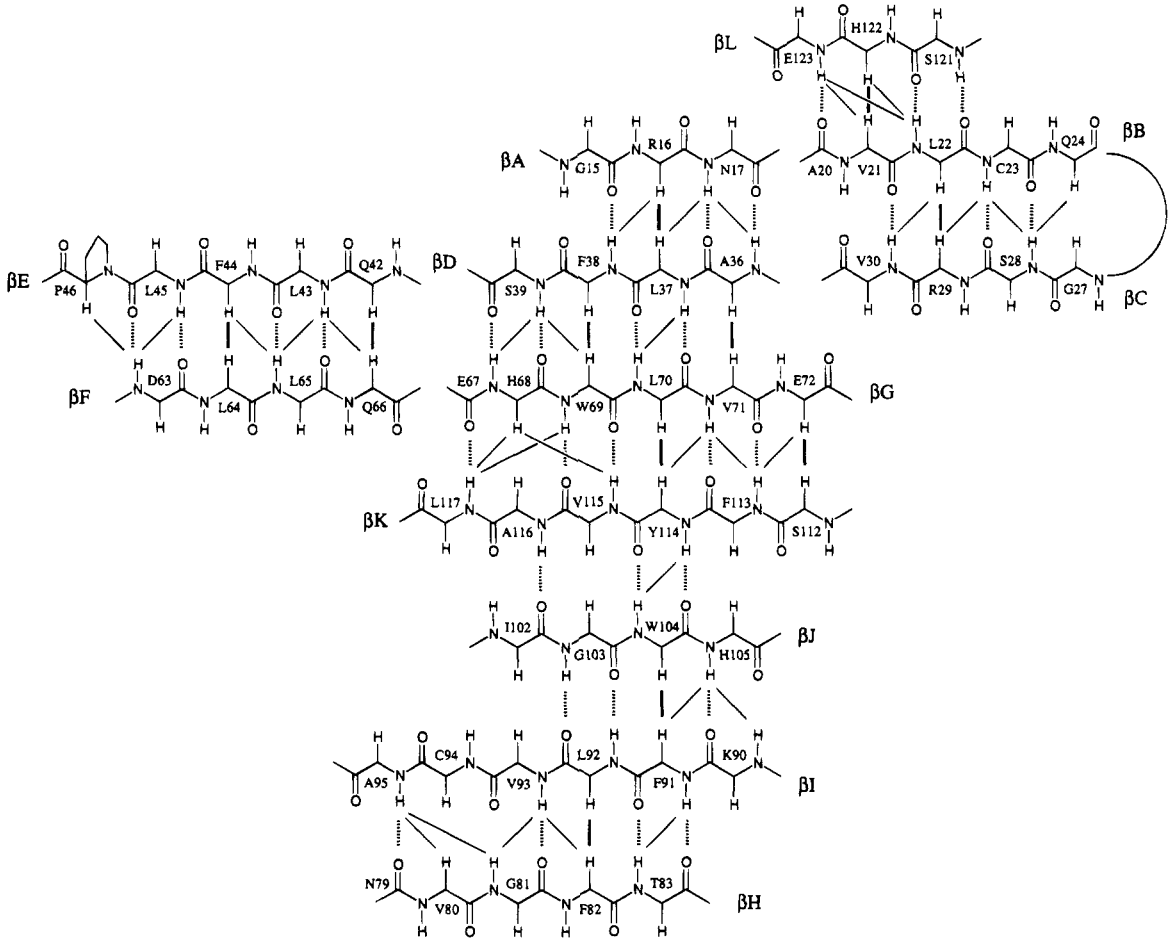

FIGURE 6: β sheets of hMss4. The strands are labeled in the order of amino acid sequence. The interstrand NOEs and hydrogen bonds are represented by solid and dashed lines, respectively.

trum demonstrated that some of these low-field C_αHs show NOEs to each other. It was based on these cross strand NOEs that we first constructed several strands of the β sheet. Other cross strand NOEs from the NHs of one strand to the C_αHs of another allowed us to extend the nascent sheet (Figure 6). Large $^3J_{NH\text{-}C_\alpha H}$ coupling constants and intense sequential $NH_i/C_\alpha H_{i-1}$ NOEs also indicate the existence of a β strand. The identification of regular structural elements was also greatly aided by the analysis of slow-exchanging amide protons that presumably form hydrogen bonds with carbonyls across the strands. Two short helical segments were also identified in a similar manner. The cross strand NOEs and hydrogen bonds are depicted in Figure 6, while the secondary structure elements found are shown in Figure 7, together with qualitative intensities of the sequential NOEs and values of the $^3J_{NH\text{-}C_\alpha H}$ coupling constants.

The secondary structure of Mss4 can be described as a seven-stranded antiparallel β sheet flanked by two smaller sheets comprising two and three antiparallel strands, respectively. In the following discussion, we have sequentially numbered the β strands and helices whereas the loops are referred to by the letters of the two adjacent secondary elements. Aside from the β strands, other regular secondary elements include two short helices: helix αA (residues 74−77) connects strands βG and βH, while helix αB (residues 117−120) is located between βK and βL. It is not clear whether these helices are alpha or 3_{10} helices at the present stage of structural analysis. The central sheet includes strands βA, βD, βG, βK, βJ, βI, and βH with a connection topology of $+1, +1, +4x, -1, -1, -1$. To our knowledge, the overall fold of hMss4 is not significantly similar to any proteins with known tertiary structures. It therefore represents a novel fold. On the basis of our NMR data, we estimate that hMss4 contains about 44% β sheet, 6% helix, and 50% loop structure. As mentioned previously, Cys 23, Cys 26, Cys 94, and Cys 97 are likely to bind Zn^{2+}, and they are located at the BC and GH loops. Many residues belonging to the EF loop did not show amide $^1H/^{15}N$ correlation peaks in the HSQC spectrum, indicating that this loop is highly mobile.

DISCUSSION

Due to their ability to exist in both GTP- and GDP-bound forms, Ras proteins function as molecular switches and thereby regulate the transmission of cellular signals (Hall, 1990; Bourne et al., 1991). However, the interconversion between the GTP- and GDP-bound forms of many Ras proteins is slow. Accessory proteins that accelerate these processes have been identified. For example, GAPs that promote the intrinsic activity of Ras proteins to hydrolyze GTP have been identified for Ras and Rho (Boguski & McCormick, 1993). Similarly, CDC25, mSOS, ras-GRF, Dbl, and RCC1 promote the exchange of GTP for GDP in Ras, Rho, or Ran (Feig, 1994). As with the Ras proteins, the family of newly identified Ras accessory proteins is growing. More recently, DSS4 and Mss4 were shown to be GEFs for the Sec4/Ypt1/Rab branch of Ras GTPases that regulate vesicular transport from yeast to mammals (Burton

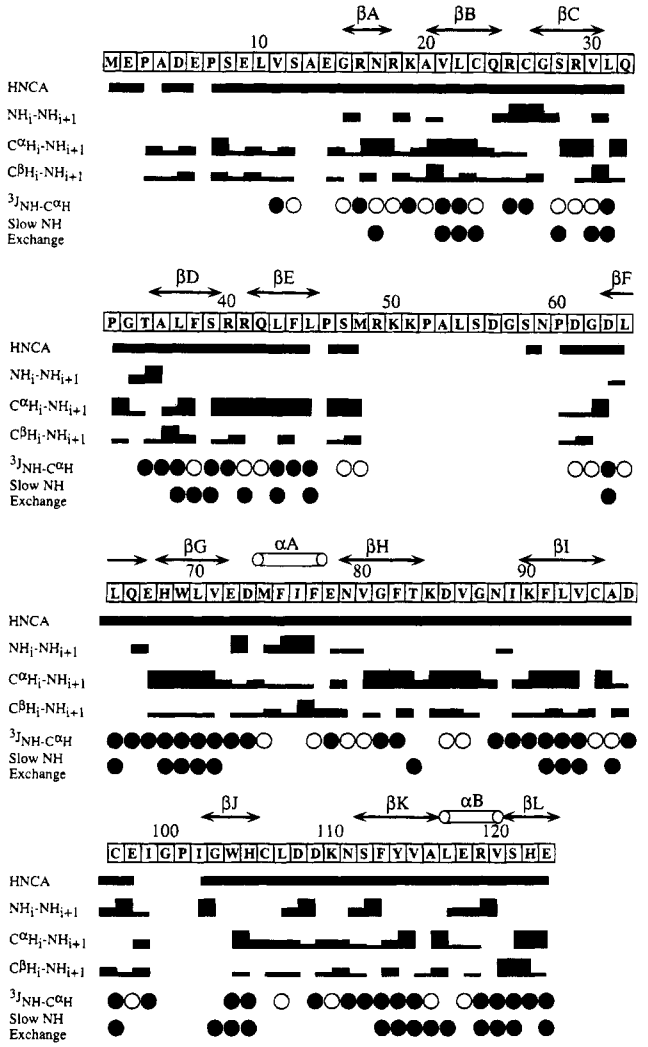

FIGURE 7: Secondary structure of hMss4 together with a summary of sequential NOEs, $^3J_{NH\text{-}C_\alpha H}$ coupling constants, and slowly exchanging amide protons. The widths of the bars indicate the intensities of the corresponding NOEs (strong, medium, or weak). Small and large $^3J_{NH\text{-}C_\alpha H}$ coupling constants are shown as open and closed circles, respectively. Slowly exchanging amide protons are also represented by closed circles. The proposed secondary structural elements are shown above the amino acid sequence.

et al., 1993; Burton et al., 1994). DSS4 and Mss4 share significant amino acid sequence homology; however, neither contains recognizable similarities with other GEFs such as CDC25, mSOS, and ras-GRF. Despite the high sequence homology among Ras family members, a particular GEF only promotes the guanine nucleotide exchange of a subset of genetically and functionally related Ras proteins. Structures of both the GDP- and GTP-bound forms of $p21^{ras}$ have been solved by X-ray crystallography (de Vos et al., 1988; Pai et al., 1989). These structures have guided the genetic and mutagenesis analysis of Ras-related proteins, including those involved in vesicular transport. For instance, by constructing Sec4/Ypt1 chimeric proteins with swapped domains, Brennwald et al. have demonstrated that Rabs are not sufficient to determine the docking specificity of transporting vesicles (Brennwald & Novick, 1993). Similarly, Sec9 has been shown to be an effector of Sec4 because it can suppress a Sec4 effector domain mutant (Brennwald et al., 1994). On the other hand, lack of structural information has severely hindered investigations on the function and specificity of guanine nucleotide exchange factors. To gain insights into the molecular recognition between Ras proteins and GEFs, and into the mechanism of guanine nucleotide exchange, we have selected the hMss4 GEF and its Ras targets as a model system. As a first step toward these goals, we have cloned, overexpressed, biochemically characterized, and sequentially assigned the human Mss4 protein.

A number of nonconserved amino acid substitutions reside in the N-terminal regions of rat and human Mss4. The ^1H and ^{15}N signals of these residues in the human protein are significantly sharper than those elsewhere in the protein. This suggests that the N-terminal region of Mss4 adopts random coil conformations in solution and, therefore, that the difference in sequence is unlikely to perturb its structure. The four cysteines presumed to coordinate the Zn^{2+} ion (Cys 23, Cys 26, Cys 94, and Cys 97) are strictly conserved between the rat and human Mss4. Although the equivalent residues in DSS4 for Cys 23 and Cys 26 cannot be easily identified on the basis of the sequence alignment, Cys 94 and Cys 97 are conserved between Mss4 and DSS4, suggesting that DSS4 might also be a Zn^{2+}-binding protein. Other GEFs, such as CDC25, mSOS, and ras-GRF, contain a homologous region of about 250 amino acids (Shou et al., 1992), yet there is no similar pattern of cysteine distribution in this region. Whether Zn^{2+} binding is a common feature among GEFs, however, remains to be seen.

Like other GEFs, Mss4 seems to promote GDP−GTP exchange by stabilizing the nucleotide-free form of its target Ras-like small GTPases (Mitsou et al., 1992; Lai et al., 1993; Feig, 1994; Burton, 1994). The regions of several Ras proteins that interact with their GEFs have been defined, and these studies indicate that the GEF-binding surfaces of various Ras proteins might be different (Mitsou et al., 1992; Segal et al., 1993; Burstein et al., 1992; Brondyk et al., 1993). So far, there is no structural information available for GEF family members that allow the identification of their binding interfaces for Ras. On the basis of the structural insights gained in this study, we speculate that the Zn^{2+}-binding site of Mss4 or DSS4 is likely to be involved in its association with interacting Ras proteins. The region surrounding Cys 94 and Cys 97 of Mss4 is highly conserved between Mss4 and DSS4, and these proteins share many common Ras proteins as targets. Interestingly, a mutant of *DSS4* that can suppress the temperature sensitivity of *sec4−8* contains an Asp to Gly mutation at a position next to one of the Zn^{2+}-binding cysteines, corresponding to Cys 97 in Mss4 (Moya et al., 1993). A large insertion also occurs immediately N-terminal to this region in the DSS4 protein. Although the large size of the complex formed by Mss4 and its substrate Ras protein poses a significant challenge for structure determination by NMR, techniques such as chemical shift perturbation should be useful in efforts to determine the interacting surfaces and to provide clues about the mechanism of Mss4. Experiments designed to confirm this hypothesis are in progress.

ACKNOWLEDGMENT

We thank Dr. Sheng Luan and Dr. Xiao-Feng Zheng for their assistance in the cloning of *mss4* and *SEC4*, Dr. Franz Bruckert and Dr. Morgen Søgaard in the Rothman Laboratory at the Memorial Sloan Kettering Cancer Center for Ypt1 protein, Dr. Shaw Huang for his assistance in NMR spectroscopy, and Larry Myers for his help with the zinc-binding assay.

SUPPORTING INFORMATION AVAILABLE

Listing of the ^{1}H, ^{15}N, and ^{13}C chemical shifts of the recombinant human Mss4 (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Bax, A., & Ikura, M. (1991) *J. Biomol. NMR 1*, 99−104.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188−7190.

Bax, A., Clore, G. M., & Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425−431.

Berg, J. M. (1995) *Acc. Chem. Res. 28*, 14−19.

Boguski, M. S., & McCormick, F. (1993) *Nature 366*, 643−654.

Bourne, H. R., Sanders, D. A., & McCormick, F. (1991) *Nature 349*, 117−127.

Brennwald, P., & Novick, P. (1993) *Nature 362*, 560−563.

Brennwald, P., Kearns, B., Champion, K., Keränen, S., Bankaitis, V., & Novick, P. (1994) *Cell 79*, 245−258.

Brondyk, W. H., McKiernan, C. J., Burstein, E. S., & Macara, I. G. (1993) *J. Biol. Chem. 268*, 9410−9415.

Burstein, E. S., Brondyk, W. H., & Macara, I. G. (1992) *J. Biol. Chem. 267*, 22715−22718.

Burton, J., Roberts, D., Montaldi, M., Novick, P., & De Camilli, P. (1993) *Nature 361*, 464−467.

Burton, J., Burns, M. E., Gatti, E., Augustine, G. L., & De Camilli, P. (1994) *EMBO J. 13*, 5547−5558.

De Vos, A. M., Tong, L., Milburn, M. V., Matias, P. M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., & Kim, S.-H. (1988) *Science 239*, 888−893.

Dunn, B., Stearns, T., & Botstein, D. (1993) *Nature 362*, 563−565.

Feig, L. A. (1994) *Curr. Opin. Cell Biol. 6*, 204−211.

Ferro-Novick, S., & Jahn, R. (1994) *Nature 370*, 191−193.

Fischer von Mollard, G., Stahl, B., Li, C., Südhof, T. C., & Jahn, R. (1994) *Trends Biochem. Sci. 19*, 164−168.

Geppert, M., Bolshakov, V. Y., Siegelbaum, S. A., Takei, K., De Camilli, P., Hammer, R. E., & Südhof, T. C. (1994) *Nature 369*, 493−497.

Giedroc, D. P., Keating, K. M., Williams, K. R., Konigsberg, W. H., & Coleman, J. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8452−8456.

Goud, B., Salminen, A., Walworth, N. C., & Novick, P. J. (1988) *Cell 53*, 753−768.

Hall, A. (1990) *Science 249*, 635−640.

Hare, D. (1991) *FELIX Manual* Version 2.0.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.

Kay, L. E., et al. (1990) *J. Magn. Reson. 89*, 496−514.

Lai, C.-C., Boguski, M., Broek, D., & Powers, S. (1993) *Mol. Cell Biol. 13*, 1345−1352.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.

Mitsou, M. Y., Jacquet, E., Poullet, P., Rensland, H., Gidion, P., Schlichting, I., Wittinghofer, A., & Parmeggiani, A. (1992) *EMBO J. 11*, 2391−2397.

Moya, M., Roberts, D., & Novick, P. (1993) *Nature 361*, 460−463.

Myers, L. C., Verdine, G. L., & Wagner, G. (1993) *Biochemistry 32*, 14089−14094.

Pai, E. F., Kabsch, W., Krengel, U., Holmes, K. C., John, J., & Wittinghofer, A. (1989) *Nature 341*, 209−214.

Qian, X., Jeon, C., Yoon, H., Agarwal, K., & Weiss, M. A. (1993) *Nature 365*, 277−279.

Segal, M., Willumsen, B. M., & Levitzki, A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5564−5568.

Shou, C., Farnsworth, C. L., Neel, B. G., & Feig, L. A. (1992) *Nature 358*, 351−354.

Novick, P. J., & Brennwald, P. (1993) *Cell 75*, 597−601.

Pfeffer, S. R. (1994) *Curr. Opin. Cell Biol. 6*, 522−526.

Rothman, J. E. (1994) *Nature 372*, 55−63.

Salminen, A., & Novick, P. J. (1987) *Cell 49*, 527−538.

Segev, N., Mulholland, J., & Botstein, D. (1988) *Cell 52*, 915−924.

Søgaard, M., Tani, K., Ye, R. R., Geromanos, S., Tempst, P., Kirchhausen, T., Rothman, J. E., & Söllner, T. (1994) *Cell 78*, 937−948.

Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.

BI950449I