A novel actin-bundling/filopodium-forming domain conserved in insulin receptor tyrosine kinase substrate p53 and missing in metastasis protein

Akiko Yamagishi, Michitaka Masuda§, Takashi Ohki, Hirofumi Onishi, and Naoki Mochizuki

Department of Structural Analysis, National Cardiovascular Center Research Institute

5-7-1 Fujishiro-dai, Suita, Osaka 565-8565, Japan

§Corresponding author: Michitaka Masuda

Phone: +81-6-6833-5012 Ext. 2508

Fax: +81-6-6835-5461

E-mail: masuda61@ri.ncvc.go.jp

Running title: IRSp53 induces filopodia by a novel actin bundling domain
SUMMARY

Insulin receptor tyrosine kinase substrate p53 (IRSp53) has been identified as an SH3 domain-containing adaptor that links Rac1 with a Wiskott-Aldrich syndrome family verprolin-homologous protein 2 (WAVE2) to induce lamellipodia or Cdc42 with mammalian enabled (Mena) to induce filopodia. The recruitment of these SH3-binding partners by IRSp53 is thought to be crucial for F-actin rearrangements. Here, we show that the N-terminal predicted helical stretch of 250 amino acids of IRSp53 is an evolutionarily conserved F-actin bundling domain involved in filopodium formation. Five proteins including IRSp53 and missing in metastasis protein (MIM) share this unique domain and are highly conserved in vertebrates. We named the conserved domain IRSp53/MIM homology domain (IMD). The IMD has domain relatives in invertebrates but does not show obvious homology to any known actin interacting proteins. The IMD alone, derived from either IRSp53 or MIM, induced filopodia in HeLa cells and the formation of tightly-packed parallel F-actin bundles in vitro. These results suggest that IRSp53 and MIM belong to a novel actin bundling protein family. Furthermore, we found that filopodium-inducing IMD activity in the full length IRSp53 was regulated by active Cdc42 and Rac1. The SH3 domain was not necessary for IMD-induced filopodium formation. Our results indicate that IRSp53, when activated by small GTPases, participates in F-actin reorganization not only in an SH3-dependent manner but also in a manner dependent on the activity of the IMD.
INTRODUCTION

Insulin receptor tyrosine kinase substrate p53 (IRSp53)\(^1\), also known as brain-specific angiogenesis inhibitor 1 associated protein 2, is a multifunctional adaptor protein enriched in the central nervous system (1-3). The protein contains a unique N-terminal 250 amino acid stretch, a half-Cdc42/Rac interactive binding (CRIB) motif, a proline-rich domain, a Src homology 3 (SH3) domain, and a WW domain-binding motif (WWB). IRSp53 is directly regulated by Rho family small GTPases, Rac1 and Cdc42, and provides a molecular link between these GTPases and the actin cytoskeleton regulators Wiskott-Aldrich syndrome protein (WASP) family verprolin homologous-protein 2 (WAVE2) and mammalian enabled (Mena), which are involved in the formation of lamellipodia (4, 5) and filopodia (6, 7). Active Cdc42 binds to the half-CRIB motif (6, 7), whereas Rac1 binds to the unique N-terminal domain (8). The association of Rac1 or Cdc42 is proposed to liberate the C-terminal SH3 domain masked intra-molecularly by its N-terminus, thereby allowing the SH3 domain to interact with its binding partners (4, 7, 9). Thus, the SH3 domain is thought to be essential for IRSp53-mediated actin reorganization. However, the N-terminal half of IRSp53 lacking the SH3 domain was reported to induce neurite outgrowth in a neuroblastoma cell line (6) and filopodia in B16 melanoma cells (10), suggesting that IRSp53 promotes actin reorganization independently of SH3-domain-mediated inter-molecular interactions.

Recently, a novel monomeric actin binding protein, missing in metastasis protein (MIM), containing a WASP homology 2 (WH2) domain in the C-terminus, was reported in human and mouse (11-13), and found to share the unique N-terminal domain with IRSp53 (13). We found
that the N-terminal domains of MIM and IRSp53 also share other characteristic features; the predicted secondary structures are almost purely helical (see (9) for IRSp53) and the estimated isoelectric points are around nine. MIM induces actin cytoskeleton reorganization in cultured cells. This activity is not dependent on the C-terminal half (13), suggesting that the N-terminal half containing the IRSp53 homologous domain plays a key role in actin reorganization.

Here we show that IRSp53 and MIM belong to an evolutionarily related protein family sharing a well-conserved N-terminal helical domain (IMD, IRSp53/MIM homology domain) as a key constituent. We investigated the role of the IMD in actin reorganization. Our results indicate that the IMDs of IRSp53 and MIM induce filopodia in cultured cells and form tightly packed F-actin bundles in vitro. The filopodium-forming activity of the IMD in full length IRSp53 is regulated by small GTPases. Thus, upon association with active Rac1 or Cdc42, IRSp53 can induce actin cytoskeleton reorganization by dual mechanisms: The SH3-mediated recruitment of F-actin regulators and the action of the novel actin bundling domain in the N-terminus. Both mechanisms may work synergistically or additively in controlling cortical actin dynamics.
EXPERIMENTAL PROCEDURES

Data base search—Proteins homologous to IRSp53 and MIM were identified in the GenBank database using BLAST on the National Center for Biotechnology Information (NCBI) Web site and on the GenomeNet Web site. Details of sequences thus retrieved are described in Supplemental data, *Collection of nucleotide sequence data of putative members of the IRSp53/MIM family proteins*. The CLUSTALW engine at the GenomeNet Web site was used to align amino acid sequences and to construct phylogenetic trees. The secondary structures of proteins were predicted by 3D-PSSM (14).

Plasmids—An IRSp53 expression vector, pEF-BOS-Myc-IRSp53 (human isoform 1) was kindly donated by Dr. Miki (4). cDNAs encoding full length IRSp53 (amino acids 1-521), IRSp53-ΔSH3 (aa 1-364), IRSp53-IMD (aa 1-250) and IRSp53-ΔIMD (aa 251-521) (Fig. 1B) were amplified by polymerase chain reaction (PCR) and inserted into pEGFP-C1 (BD Bioscience Clontech, Palo Alto, CA), pCXN2-FLAG (15), and pGEX-4T3 or 6P3 (Amersham Biosciences, Piscataway, NJ) vectors. The DNA fragments encoding IRSp53 where Arg was substituted for Trp$^{413}$ or Ala for both Phe$^{427}$ and Pro$^{428}$ in the SH3 domain, hereafter referred to as IRSp53-W/R or IRSp53-FP/AA, were amplified by PCR and ligated into pEGFP-C1. cDNA of Rac1V12, Rac1N17, Cdc42V12, or Cdc42 N17 was subcloned into pIRM21, an expression vector expressing FLAG-tagged protein and internal-ribosomal-entry-site-driven dsFP593 (16). A cDNA clone encoding N-terminal fragment (aa 1-430) of human MIM was obtained by PCR from a human brain cDNA library (BD Bioscience Clontech). The cDNA encoding C-terminal MIM (aa 400-755, KIAA0429) was obtained from Kazusa DNA Research Institute. The full
length MIM cDNA was amplified through overlap PCR using these N- and C-terminal cDNAs as templates. The cDNAs encoding the full length human MIM (aa 1-755), MIM IMD (aa 1-250) and MIM-∆IMD (aa 251-755) were inserted into pEGFP-C1, pCXN2-FLAG and pGEX-4T3 or 6P3 vectors.

**Cells and transfection**—HeLa cells and 293T cells were cultured in Dulbecco’s modified Eagles medium containing 10% fetal bovine serum and 2 mM L-glutamine. HeLa cells and 293T cells were transfected using LipofectAMINE2000 (Invitrogen Corp., Carlsbad, CA) according to the manufacturer’s protocol.

**Antibodies and immunofluorescence analysis**—Rhodamine-conjugated phalloidin, and Alexa546-conjugated anti-mouse IgG were purchased from Molecular Probes (Eugene, OR); anti-FLAG M2 antibody was from Sigma-Aldrich (St. Louis, MO). HeLa cells transfected with the plasmids indicated in the figures and cultured for 15-18 h were fixed with 2% formaldehyde in phosphate buffered saline (PBS) and permeabilized with 0.1% Triton X-100 in PBS. Cells transfected with plasmids expressing green fluorescent protein (GFP)-tagged proteins were counter-stained with Rhodamine-phalloidin. Cells transfected with both GFP-tagged protein-expressing vectors and FLAG-tagged small GTPase-expressing vectors were immunostained with anti-FLAG M2 antibody followed by an Alexa546-conjugated anti-mouse IgG. Fluorescence images were obtained using a confocal microscope (BX50WI, Fluoview, Olympus, Tokyo, Japan) with a water immersion objective lens (LUMPlanF1 60×/0.90 w). To show the entire cell morphology in detail, all cell images shown were extended focus images reconstructed from a series of optical sections taken at 0.2 to 0.3 μm intervals.
F-actin binding and bundling assays—F-actin was prepared from rabbit skeletal muscle as described (17). Glutathione S-transferase (GST)-fusion proteins of various fragments of IRSp53 and MIM (Fig. 1B) were expressed in BL21-Star (DE3) cells (Invitrogen Corp.), purified using glutathione-Sepharose (Amersham Biosciences), then buffer-exchanged into F-buffer (25 mM Hepes, pH 7.5, 100 mM KCl, 0.2 mM CaCl₂, 2 mM MgCl₂, 2 mM EGTA, 0.2 mM ATP, 1 mM dithiothreitol) containing 0.1% C₁₂E₈ (Nikko Chemicals, Tokyo, Japan). For binding assays, purified GST-fused fragments were clarified by centrifugation at 400,000 × g for 15 min to remove any aggregates, mixed with F-actin in the F-buffer, and incubated for 30 min on ice. The final concentration of the GST-fusions and F-actin were 1.2 µM and 5 µM (as for G-actin), respectively. The mixture was then centrifuged as above and equal aliquots of the supernatant and the pellet were analyzed by SDS-PAGE followed by Coomassie Blue staining. For quantitative analysis of F-actin binding and bundling, the IMDs were cleaved out from the GST-fusions expressed by pGEX-6P3 vectors using PreScission Protease (Amersham Biosciences) and further purified by cation-exchange chromatography (Resource S, Amersham Biosciences). To quantify F-actin binding, increasing amounts of F-actin was incubated with 2 µM IRSp53-IMD or MIM-IMD in the F-buffer for 3 h at room temperature. The samples were then centrifuged and analyzed as above. Protein bands were quantified by densitometry (Personal Densitometer SI, Amersham Biosciences). For quantitative bundling assay, increasing amounts of the IMDs were incubated with 1 µM F-actin in the F-buffer for 1 h at room temperature. The supernatant and the pellet were separated by low-speed centrifugation (10,000 × g for 30 min) and analyzed as above.
Observation of actin bundles—For fluorescence microscope observation, a fixed concentration of F-actin (1.2 µM final concentration) was mixed with variable concentrations of the GST-fused fragments (0.24 to 12 µM). After incubation for 30 min on ice in F-buffer, F-actin was stained with Rhodamine-phalloidin for 15 min on ice. The mixtures were applied to poly-L-lysine-coated glass coverslips, and incubated for 20 min at room temperature. The adherent material was washed with F-buffer and observed with the confocal laser scanning microscope using an oil immersion objective lens (PlanApo 60×/1.40 oil). For negative staining of actin filaments and bundles, the Rhodamine-phalloidin stained specimens described above were diluted ten times with F-buffer, placed onto a carbon-coated mesh and stained with 2% uranyl acetate. For observation of thin sectioned specimens, actin bundles formed after incubation for 1 h on ice were packed by centrifugation and fixed in 2.5% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.4, then sequentially incubated with 0.1% aqueous tannic acid and 0.2% uranyl acetate (18), postfixed in 0.5% aqueous OsO₄, dehydrated and embedded in Epon 812. Thin sections, stained by lead citrate, were examined with a CM 120 electron microscope (Philips Electronics, Eindhoven, The Netherlands) equipped with a multiscan cooled charge-coupled device camera (Model 791, Gatan, Pleasanton, CA).

Cross-linking of proteins—One µM of purified IRSp53-IMD, MIM-IMD and chymotrypsinogen A (Amersham Biosciences) as the control were cross-linked in 0.1 M 2-morpholinoethanesulfonic acid (pH 5.0) at room temperature. The reaction was started by the addition of 4 mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and was stopped by the addition of 50 mM Tris-HCl (pH 8.0) at the time points indicated in Fig. 7A.
**Immunoprecipitation**—293T cells were washed with PBS and lysed in lysis buffer (100 mM NaCl, 25 mM Hepes, pH 7.5, 2 mM MgCl$_2$, 2 mM EGTA, 0.5% Triton X-100, containing protease inhibitor cocktail; Roche Applied Science, Basel, Switzerland). Lysates were precleared by centrifugation at 100,000 $\times g$ for 10 min, followed by immunoprecipitation with a rabbit anti-GFP antibody and Protein A Sepharose beads (Amersham Biosciences). Immunoprecipitates were subjected to SDS-PAGE and immunoblotting with antibodies as indicated in Fig. 7B. Proteins reacting with primary antibodies were visualized by an enhanced chemiluminescence system (Amersham Biosciences) for detecting peroxidase-conjugated secondary antibodies and analyzed with an LAS-1000 system (Fuji Film, Tokyo, Japan).
RESULTS

The N-terminal helical domain is evolutionarily conserved in IRSp53 family proteins and MIM family proteins—IRSp53 and MIM share the N-terminal stretch of 250 amino acids (22% identical/18% similar) while the remaining parts of the molecules show only marginal similarity. To explore whether this similarity is based on real homology, we searched the GenBank database for proteins having similar sequences. First we found three more genes encoding homologous N-terminal sequences in the human genome: Insulin receptor tyrosine kinase substrate (IRTKS), a hypothetical gene FLJ22582 and ABBA-1 (Fig. 1A; see Supplemental Table 1 for detail). IRTKS and FLJ22582 are IRSp53-related proteins containing an SH3 domain in the C-terminal half. However, both of them lack the half-CRIB motif found in IRSp53 and FLJ22582 further lacks the WWB (PPPXY) (Fig. 1B). ABBA-1 is a MIM-related protein which possesses a WH2 domain in the C-terminus. Further database searches have shown that each of these five proteins has a putative ortholog in chicken and zebra fish, indicating that they are well conserved through vertebrate evolution (Supplemental Figs. 1-4). As pointed out previously (9, 13), related proteins are also found in invertebrates (C. elegans M04F3.5 protein and D. melanogaster CG32082 protein). In an amino acid sequence alignment of the N-terminal region of these proteins (Fig. 1A), clusters of basic amino acids, proline, glycine and clusters of hydrophobic amino acids are well conserved. There is a signature sequence of ALxEE[RK][RG]RFCx(0,1)F[IL] in the C-terminal half of the stretch. As expected from the number of basic amino acid clusters in this domain, the estimated isoelectric points are highly basic, ranging from 8.5 for human MIM to 9.2 for human FLJ22582.
The identity of the N-terminal domain is further supported by predicted secondary structures. The domains are almost purely helical, 82% to 87% for the IRSp53-related proteins, 96 to 100% for MIM-related proteins and intermediate contents of 89% and 87% for M04F3.5 and CG32082, respectively. Helix-breaking amino acid residues at the four breaking sites of the IRSp53-related proteins are also conserved in MIM/ABBA family proteins (asterisks in Fig. 1A). Thus, all of these proteins appear to have a common segmentation pattern of helices, helix I to V (Fig. 1A). Although human IRSp53 lacks helix V, it is predicted to be present in the chicken ortholog.

A phylogenetic tree (Fig. 1C) based on the alignment of the IMDs shows that the vertebrate IRSp53/MIM family is divided into two major groups: The IRSp53 sub-family and the MIM/ABBA sub-family. The putative invertebrate homologs are positioned between them. The tree of the IMDs exactly reflects the hierarchy of domain composition of these proteins. The IRSp53 sub-family members contain an SH3 domain and the MIM/ABBA sub-family proteins contain a WH2 domain. The vertebrate SH3-containing sub-family is further divided into three groups according to the presence or absence of the WWB and the half-CRIB motif. These data suggest that the IRSp53/MIM family originated from a common ancestor and diverged through evolution. This hypothesis is supported by the fact that IRTKS and FLJ22582 but not M04F3.5 or CG32082 share highly homologous C-termini with the MIM/ABBA sub-family members (Supplemental Fig. 6). Our analyses suggest the presence of an evolutionarily conserved IRSp53/MIM family and that the IMDs are the key components for the functional roles of proteins belonging to this family.
The IMDs of IRSp53 and MIM induce filopodia in HeLa cells—To explore the functional roles of the IMD, we first examined the morphological effects of ectopic expression of the IRSp53-IMD and the MIM-IMD in HeLa cells. Cells expressing the GFP-tagged IMD of IRSp53 formed numerous long filopodia that were F-actin rich as demonstrated by Rhodamine-phalloidin staining (Fig. 2, a, a’ and b, b’). The MIM-IMD also induced filopodia but these were reduced in length (Fig. 2, c, c’ and d, d’). In addition, MIM-IMD promoted the formation of microvillus-like protrusions on the apical cell surface. IRSp53-IMD and MIM-IMD localized to and occasionally were concentrated in these protrusions (arrows in Fig. 2, c’ and d’). Both IMDS appeared not to be associated with stress fibers. There were no obvious signs of enhanced lamellipodial activity or disruption of stress fibers in these IMD-expressing cells. GFP used as a negative control did not induce any morphological changes (Fig. 2, e and f). Truncated fragments of IMD, IRSp53-N-IMD (aa 1-161) and IRSp53-C-IMD (aa 105-250) could not stimulate filopodium formation (data not shown). These data indicate that both IMDS are capable of inducing filopodia in cells. Since IRSp53 and MIM represent the most divergent members of the vertebrate IRSp53/MIM protein family (Fig. 1C), the filopodium-inducing activity of the IMD is likely to be conserved in all family members.

IMD does not act upstream of Rac1 or Cdc42 for filopodium formation—Actin cytoskeletal reorganization is often a hallmark of Rho family GTPases. Previous reports have shown that Rac1 binds to the N-terminus of IRSp53 (8) and Cdc42 binds to the aa 202-305 fragment
containing the half-CRIB motif (6). Therefore, we examined whether Cdc42 or Rac1 activation was involved in IMD-induced filopodium formation. The formation of numerous filopodia induced by IRSp53-IMD was not perturbed by the co-expression of dominant negative Cdc42 or Rac1 (Fig. 3). There was no quantitative difference in the ratio of filopodium forming cells among HeLa cells transfected with IRSp53-IMD alone, those transfected with IRSp53-IMD and Cdc42N17, and those transfected with IRSp53 and Rac1N17 (Fig. 3, d), suggesting that the IMD itself is not regulated by these small GTPases. This result also excludes the possibility that the domain functions upstream of these small GTPases.

_The filopodium-inducing IMD activity of wild-type IRSp53 is regulated by Cdc42 and Rac1_—The common mechanism of effector activation by Rho family GTPases appears to be dependent on the disruption of intra-molecular autoinhibitory interactions. Cdc42-induced conformational changes have also been demonstrated for the molecule containing a half or semi-CRIB motif, Par6 (19). First we found that GFP-tagged IRSp53-WT or GFP-tagged IRSp53-ΔSH3, when expressed in moderate levels, could not induce filopodia (Fig. 4A, a, b, and c). As reported earlier (7, 10), cells expressing very high amounts of IRSp53 often formed dendritic extensions accompanied with severe retraction of the cell body. As noted in Fig. 2 legend, these cells were omitted from our analyses. Next, we examined whether the IMD function was regulated by Cdc42 and Rac1 in IRSp53-WT and in IRSp53-ΔSH3 containing the half-CRIB motif. Co-expression of the active Cdc42 with these IRSp53 constructs led to massive formation of wavy filopodia (IRSp53+Cdc42 phenotype as shown in Fig. 4A, d and e) that was clearly
distinguishable from straight filopodia induced in cells co-expressing GFP and active Cdc42 (Cdc42 phenotype as shown in Fig. 4A, f). A similar level of filopodium induction, mixed with Rac1-dependent enhanced lamellipodia activity (Fig. 4A, i) was induced by the co-expression of active Rac1 (Fig. 4A, g and h). These results suggest that the SH3 domain is not necessary for IMD-dependent filopodium formation. Our results also suggest that the filopodium-inducing IMD activity in wild-type IRSp53 is regulated by Cdc42 and Rac1. The central region of IRSp53 containing the half-CRIB motif appears to be essential for this regulation, as previously suggested for the regulation of the SH3 domain (7, 9).

To further confirm that the IMD-induced filopodium formation is independent of SH3-binding molecules, we used two non-functional SH3 mutants, IRSp53-W/R and IRSp53-FP/AA (7). Both mutants could induce filopodia when expressed with active Cdc42 (Fig. 4B). Although active Cdc42 alone induced filopodium formation in the majority of cells (Fig. 4B-d, the blank segment in the stacked bar graph reflecting the Cdc42 phenotype), exaggerated filopodium formation, an indication of IMD activity, occurred only when Cdc42 was co-expressed with IRSp53-WT and IRSp53-SH3 mutants (Fig. 4B-d, the filled segment in the stacked bar graph reflecting the IRSp53+Cdc42 phenotype). Thus, IRSp53 can promote filopodium formation independently of SH3-mediated inter-molecular interactions.

**In vitro F-actin bundling activity of IMD**—The filopodium-promoting activity of the IMDs of IRSp53 and MIM in cultured cells led us to examine whether these IMDs have F-actin binding and bundling activity. We examined F-actin binding/bundling activity of the GST-fused IMD and
other fragments, and also tag-free purified IMDs in vitro. As shown in Fig. 5A, GST-fused IRSp53-IMD, IRSp53-ΔSH3 and MIM-IMD, but not GST were co-sedimented with F-actin in a high-speed assay (total binding). To exclude the possible contribution of GST-tag or contaminating bacterial proteins to F-actin binding and bundling, the activities of purified tag-free IMDs (Fig. 5B, left panel) were examined. In the high-speed assays, the IMDs of IRSp53 and MIM bound to F-actin in a concentration-dependent and saturable manner (Fig. 5B, right panel). The apparent half-maximum concentrations of F-actin for IMD binding were almost the same, 0.5 µM, irrespective of the variation between the maximum extents of these IMDs, suggesting that both IMDs have roughly the same affinity to F-actin. Low levels of the maximum extent of bound IMDs, about 30% for IRSp53-IMD and 20% for MIM-IMD, can be explained by improper protein folding of the bacterial-made IMDs or their denaturation during the purification process.

The GST-fusions capable of F-actin binding induced thick F-actin bundles (Fig. 5C). Although GFP-tagged IRSp53-ΔSH3 required activation by Rac1 or Cdc42 for filopodium formation, GST-fused IRSp53-ΔSH3 alone could induce F-actin bundling. It is possible that the bacterial-made protein may not be folded properly to form the self-inhibitory conformation. IRSp53-ΔSH3 showed stronger bundling than IRSp53-IMD or MIM-IMD, however, different levels of bundling activity among these proteins may simply reflect differences in stability of these fusion proteins. To quantify the bundling activity, the tag-free IMDs were examined in low-speed sedimentation assay (Fig. 5 D). The bundling activity was concentration-dependent and most of F-actin could be incorporated into bundles in high concentrations of the IMDs.
The IMD-induced F-actin bundles could be seen under a phase contrast microscope and their thickness was measured at 0.1 to 0.2 μm by electron microscope observation of negatively stained materials (Fig. 6A). Observation of thin sections of the bundles revealed tight packing of parallel actin filaments in the bundles (Fig. 6B). The bundle as a whole was not a paracrystal in which actin filaments were packed into a hexagonal array with a constant spacing of 11.5 nm, as previously described (20). However, actin filaments in the bundles tended to be arranged in a line and partly packed into a hexagonal pattern (Fig. 6B, inset). The center-to-center distance between neighboring actin filaments aligned in a line was nearly constant and was measured at 11.2 nm in transverse sections (Fig. 6C). These observations indicate that IRSp53 acts as a typical parallel-actin-bundle-forming molecule such as fimbrin and fascin and suggest that the IRSp53/MIM family is a novel actin bundling protein family.

Self-association of IMD—Actin bundling activity requires at least two independent F-actin binding sites or a combination of one binding site plus a self-association site in an actin bundling domain. The ability of IMDs to form dimers or oligomers was examined by chemical cross-linking using a zero-length cross-linker EDC. The apparent molecular weight of the tag-free IMDs in SDS-PAGE was progressively shifted from 30 kDa, which matched to the calculated molecular weights of the IRS-IMD (28,972) and MIM-IMD (28,640) including a short linker sequence, to the dimer one of 60 kDa (Fig. 7A). Both of the tag-free IMDs were effectively cross-linked into dimers but not into trimers or tetramers. This result suggests that the purified IMDs are present as dimers. Next we examined the IMD’s self-association in cultured
cells. The IMD of IRSp53 or MIM could associate with each other and with the full-length molecule but not with the C-terminal half lacking the IMD (ΔIMD) in co-transfected 293T cells (Fig. 7B). Consistent with the IMD-dependent self-association, ΔIMD was not co-immunoprecipitated in any combinations whereas the full length IRSp53 and MIM associated with themselves. These results indicate that the IMD is a self-associating domain and suggest that IRSp53 and MIM can be present as dimeric forms in mammalian cells.
DISCUSSION

In this study we show that the N-terminal helical domain, the IMD, identified in IRSp53 and MIM, induces filopodium formation \textit{in vivo} and F-actin bundling \textit{in vitro} and suggests that these domains are conserved in an evolutionarily related protein family, the IRSp53/MIM family. We propose that the IRSp53/MIM family is a novel F-actin bundling protein family which includes invertebrate relatives. Although the family members are largely diverged, each vertebrate member (IRSp53, IRTKS, FLJ22582, MIM and ABBA) is highly conserved throughout the entire molecule in species ranging from fish to human. They are likely to have a common fundamental function, actin bundle formation, with different mechanisms of regulation.

Parallel actin bundles form the core structure of cellular protrusions such as filopodia, microvilli and microspikes. These structures are tightly packed, non-contractile bundles cross-linked by a class of F-actin bundling proteins, such as fascin and fimbrin, that determine an approximately 12 nm spacing between actin filaments (20, 21), and the involvement of such proteins is essential for structure formation (22-26). The other class of bundling proteins, represented by \(\alpha\)-actinin, are found in contractile bundles such as stress fibers. There, they cross-link actin filaments with a wide spacing of about 36 nm (27, 28) which allows myosin II to interact with the actin filaments (29). We have shown that the N-terminus of IRSp53 induces \textit{in vitro} formation of tightly packed F-actin bundles of 11 nm spacing. The localization of the N-terminal helical domain of IRSp53 and MIM in filopodia with F-actin but not in stress fibers is consistent with the idea that the protein functions in cells as a parallel F-actin bundling protein.

Our present study indicates that SH3-mediated interactions are not always necessary for
IRSp53-induced filopodium formation and this is consistent with a recent report showing that Mena and vasodilator-stimulated phosphoprotein (VASP) are not essential for this process in B16 melanoma cells (10). However, our results neither rule out the Rac1-IRSp53-WAVE2 or Cdc42-IRSp53-Mena pathway nor exclude any contribution of the C-terminal half to IRSp53 induced F-actin rearrangements. The SH3-mediated interactions could contribute to IRSp53 functions by two possible mechanisms. In the first, as in the classical view of IRSp53 function, SH3-ligands play a crucial role in actin cytoskeleton dynamics, which may additively or synergistically work with the N-terminal IMD. Among these ligands, enabled (Ena) /VASP family proteins have been reported to have actin bundling activity associated with the Ena/VASP homology 2 domain (30). Recently IRSp53 has been shown to bind to neural isoforms of espin (31), a novel parallel actin bundling protein originally identified as a component of the Sertoli-cell-spermatid ectoplasmic specialization (32). The resultant multi-domain actin bundling protein complexes may bundle F-actin with increased efficiency or contribute to changes in F-actin dynamics. Secondly, the SH3-mediated interaction could determine the localization of IRSp53. Although we and others (10) have shown that IRSp53 is able to self-localize in filopodia using its N-terminus, levels of accumulation appear not to be high. Considering that actin bundling proteins require a relatively high molar ratio to actin in order to function, this level of specificity may not be sufficient to support dynamic behavior of the cell periphery in non-transfected cells. Both WAVE2 and Mena are shown to localize at the filopodial tip (33-35), again suggesting the functional redundancy of these protein complexes with increased specificity of localization.
Here we show that the activity of the IMD is tightly regulated by Rac1 and Cdc42, in a similar manner to that of the SH3 domain (4, 7). Our results suggest that the central region of IRSp53, including the half-CRIB motif, is essential for the autoinhibition of the IMD. The N-terminus aa 1-178 of IRSp53 has been shown to interact with the region around the half-CRIB motif and inhibit binding of the SH3 domain to Mena (7). The autoinhibitory mechanisms of the IMD and the SH3 domain may work together within the same molecule. Conversely, F-actin association of the IMD and the SH3-ligand binding are likely to activate or stabilize each other.

We propose that IRSp53 is a direct effector of Cdc42 and Rac1, acting in concert with various partner proteins recruited by the SH3 domain. Further analyses are required to evaluate the activities of various IRSp53/partner protein complexes and their specific roles in the regulation of cortical actin dynamics. Although MIM has been shown to interact with protein tyrosine phosphatase delta (13), its regulation remains unknown. Our present study reveals that the IMDs are highly conserved both structurally and functionally. So far we have not found any apparent sequence homology of this domain with known F-actin interacting proteins. Future work including crystallographic studies will be needed to ascertain precise molecular mechanisms for F-actin bundling by the IMDs as well as to clarify their regulation, especially by small GTPases in IRSp53.
REFERENCES

1. Abbott, M. A., Wells, D. G., and Fallon, J. R. (1999) J.Neurosci. 19, 7300-7308

2. Oda, K., Shiratsuchi, T., Nishimori, H., Inazawa, J., Yoshikawa, H., Taketani, Y., Nakamura, Y., and Tokino, T. (1999) Cytogenet.Cell Genet. 84, 75-82

3. Yeh, T. C., Ogawa, W., Danielsen, A. G., and Roth, R. A. (1996) J.Biol.Chem. 271, 2921-2928

4. Miki, H., Yamaguchi, H., Suetsugu, S., and Takenawa, T. (2000) Nature 408, 732-735

5. Takenawa, T. and Miki, H. (2001) J.Cell Sci. 114, 1801-1809

6. Govind, S., Kozma, R., Monfries, C., Lim, L., and Ahmed, S. (2001) J.Cell Biol. 152, 579-594

7. Krugmann, S., Jordens, I., Gevaert, K., Driessens, M., Vandekerckhove, J., and Hall, A. (2001) Curr.Biol. 11, 1645-1655

8. Miki, H. and Takenawa, T. (2002) Biochem.Biophys.Res.Commun. 293, 93-99

9. Alvarez, C. E., Sutcliffe, J. G., and Thomas, E. A. (2002) J.Biol.Chem. 277, 24728-24734

10. Nakagawa, H., Miki, H., Nozumi, M., Takenawa, T., Miyamoto, S., Wehland, J., and Small, J. V. (2003) J.Cell Sci. 116, 2577-2583

11. Lee, Y. G., Macoska, J. A., Korenchuk, S., and Pienta, K. J. (2002) Neoplasia. 4, 291-294

12. Mattila, P. K., Salminen, M., Yamashiro, T., and Lappalainen, P. (2003) J.Biol.Chem. 278, 8452-8459

13. Woodings, J. A., Sharp, S. J., and Machesky, L. M. (2003) Biochem.J. 371, 463-471

14. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) J.Mol.Biol. 299, 499-520

15. Niwa, H., Yamamura, K., and Miyazaki, J. (1991) Gene 108, 193-199

16. Nagashima, K., Endo, A., Ogita, H., Kawana, A., Yamagishi, A., Kitabatake, A., Matsuda, M., and Mochizuki, N. (2002) Mol.Biol.Cell 13, 4231-4242

17. Spudich, J. A. and Watt, S. (1971) J.Biol.Chem. 246, 4866-4871

18. Svitkina, T. M., Bulanova, E. A., Chaga, O. Y., Vignjevic, D. M., Kojima, S., Vasiliev, J. M., and Borisy, G. G. (2003) J.Cell Biol. 160, 409-421
19. Garrard, S. M., Capaldo, C. T., Gao, L., Rosen, M. K., Macara, I. G., and Tomchick, D. R. (2003) *EMBO J.* **22**, 1125-1133

20. DeRosier, D. J. and Tilney, L. G. (1982) *Cold Spring Harb.Symp.Quant.Biol.* **46 Pt 2**, 525-540

21. Volkmann, N., DeRosier, D., Matsudaira, P., and Hanein, D. (2001) *J.Cell Biol.* **153**, 947-956

22. Bartles, J. R. (2000) *Curr.Opin.Cell Biol.* **12**, 72-78

23. DeRosier, D. J. and Tilney, L. G. (2000) *J.Cell Biol.* **148**, 1-6

24. Small, J. V., Stradal, T., Vignal, E., and Rottner, K. (2002) *Trends Cell Biol.* **12**, 112-120

25. Svitkina, T. M., Bulanova, E. A., Chaga, O. Y., Vignjevic, D. M., Kojima, S., Vasiliev, J. M., and Borisy, G. G. (2003) *J.Cell Biol.* **160**, 409-421

26. Vignjevic, D., Yarar, D., Welch, M. D., Peloquin, J., Svitkina, T., and Borisy, G. G. (2003) *J.Cell Biol.* **160**, 951-962

27. Meyer, R. K. and Aebi, U. (1990) *J.Cell Biol.* **110**, 2013-2024

28. Taylor, K. A., Taylor, D. W., and Schachat, F. (2000) *J.Cell Biol.* **149**, 635-646

29. Katoh, K., Kano, Y., Masuda, M., Onishi, H., and Fujiwara, K. (1998) *Mol.Biol.Cell* **9**, 1919-1938

30. Bachmann, C., Fischer, L., Walter, U., and Reinhard, M. (1999) *J.Biol.Chem.* **274**, 23549-23557

31. Sekerkova, G., Loomis, P. A., Changyaleket, B., Zheng, L., Eytan, R., Chen, B., Mugnaini, E., and Bartles, J. R. (2003) *J.Neurosci.* **23**, 1310-1319

32. Bartles, J. R., Wierda, A., and Zheng, L. (1996) *J.Cell Sci.* **109 (Pt 6)**, 1229-1239

33. Bear, J. E., Svitkina, T. M., Krause, M., Schafer, D. A., Loureiro, J. J., Strasser, G. A., Maly, I. V., Chaga, O. Y., Cooper, J. A., Borisy, G. G., and Gertler, F. B. (2002) *Cell* **109**, 509-521

34. Krause, M., Bear, J. E., Loureiro, J. J., and Gertler, F. B. (2002) *J.Cell Sci.* **115**, 4721-4726

35. Nozumi, M., Nakagawa, H., Miki, H., Takenawa, T., and Miyamoto, S. (2003) *J.Cell Sci.* **116**, 239-246
Acknowledgments: We thank Hiroaki Miki (Institute of Medical Science, Tokyo University, Tokyo) for pEF-BOS-Myc-IRSp53; Masahiko Hibi (Riken Center for Developmental Biology) for Zebrafish cDNA; Michiyuki Matsuda (Research Institute Microbial Diseases, Osaka University, Suita) for helpful discussion; Manami Sone and Eiko Moriishi for technical assistance; Hitomi Shimamoto for preparing the manuscript.
This study was supported in part by grants from the Ministry of Health, Labour and Welfare, grants from the Organization for Pharmaceutical Safety and Research of Japan, Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology, and from the Human Science Foundation of Japan.

1 The abbreviations used are: IRSp53, insulin receptor tyrosine kinase substrate p53; CRIB, Cdc42/Rac interactive binding; SH3, Src homology 3; WWB, WW domain-binding motif; WASP, Wiskott-Aldrich syndrome protein; WAVE2, WASP family verprolin-homologous protein 2; Mena, mammalian enabled; MIM, missing in metastasis protein; WH2, WASP homology 2; IMD, IRSp53/MIM homology domain; PCR, polymerase chain reaction; PBS, phosphate buffered saline; GFP, green fluorescent protein; GST, glutathione S-transferase; PAGE, polyacrylamide gel electrophoresis; EDC, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide; IRTKS, insulin receptor tyrosine kinase substrate; VASP, vasodilator-stimulated phosphoprotein; Ena, enabled.

Human ABBA-1: GenBank = GenBank Accession Number AB115770
FIGURE LEGENDS

Fig. 1. The N-terminal helical domains of IRSp53 and MIM are evolutionarily conserved.
A. Amino acid sequence alignment of the N-terminus of known human proteins containing the IRSp53/MIM homology domain (IMD). Areas where all the residues are identical are shown in black; highly conserved, dark grey; weakly conserved, light grey. Numbered lines on the sequence show the predicted helical stretches of IRSp53. Asterisks show conserved helix-breaking residues.

B. Schematic representation of the domain composition of human proteins containing the IMD. IRSp53 contains a half-CRIB motif (CRIB), an SH3 domain (SH3) and a WW domain-binding motif (WWB). IRTKS lacks CRIB while FLJ22582 lacks both CRIB and WWB. MIM and ABBA-1 each contains a WH2 domain in the C-terminus. Truncated fragments of IRSp53 and MIM used in this study are also shown. The positions of mutation in the SH3 domain are indicated by asterisks.

C. Unrooted tree of IMDs based on ClustalW alignment. Species used: human (Hs), chicken (Gg), zebra fish (Dr), D. melanogaster (Dm), and C. elegans (Ce). Members containing a WWB and a half-CRIB motif are encircled by a dotted circle and a continuous circle, respectively. See Supplemental Fig. 6 for a full version of the alignment which was further modified to improve the alignment in the N-termini and to match with the predicted helical structures.

Fig. 2. The IMDs of IRSp53 and MIM induce filopodia.

HeLa cells were transfected with the GFP-tagged IMD of IRSp53 (a, b), that of MIM (c, d) or
GFP alone (e, f). The GFP signal (left column) and F-actin visualized with Rhodamine-phalloidin (right column) are shown. a’, b’, c’ and d’ are enlarged micrographs of rectangular areas of the corresponding figures. Arrows in c’ and d’ show accumulation of MIM-IMD in filopodium/microvillus-like protrusions. Scale bar, 20 µm. All analyses of cellular phenotypes in this study were based on observations of cells expressing moderate levels of GFP-tagged IRSp53, MIM or their fragments. High-expressers often showed a dendritic phenotype with severe retraction of the cell body, and thus were not used for the analyses.

Fig. 3. The IRSp53-IMD induces filopodia independently of small GTPases.
HeLa cells expressing the GFP-fused IRSp53-IMD (a, b, c). Cells were co-transfected with FLAG-tagged dominant negative Cdc42N17 (b) or Rac1N17 (c). The upper row shows GFP signals and the lower row shows Cdc42-N17 (b) or Rac1N17 (c) expression detected by FLAG-tag. Scale bar, 20 µm. The bar graph in d shows the frequency of cells developing numerous filopodia in a typical experiment. More than 100 cells were analyzed for each data point. Cells expressing both GFP and FLAG-tagged proteins were analyzed.

Fig. 4. The IMD in wild-type IRSp53 is regulated by Cdc42 and Rac1.
A. GFP signals of HeLa cells expressing the GFP-tagged full length IRSp53 (WT, a, d, g), the GFP-tagged N-terminal half (∆SH3, b, e, h) or GFP alone (GFP, c, f, i). a, b, c; cells without co-transfection. d, e, f; cells co-transfected with Cdc42V12. g, h, i; cells co-transfected with Rac1V12. Enlarged views of the rectangular areas are shown below. Scale bar, 20 µm.
B. HeLa cells expressing the GFP-tagged full length IRSp53 (WT, a), an SH3 mutant with W417R substitution (W/R, b) or another SH3 mutant with F427A/P428A substitutions (FP/AA, c). The upper row is without co-transfection and the lower row shows cells co-transfected with the FLAG-tagged active Cdc42. The expression of Cdc42 was detected by FLAG immunostaining (not shown). Scale bar, 20 µm. The stacked bar graph in d shows the frequency of cells presenting the IRSp53+Cdc42 phenotype (filled) and the Cdc42 phenotype (blank) in a typical experiment. Cells forming numerous, long, wavy and often branching filopodia (A-d for an example) were counted as the IRSp53+Cdc42 phenotype cells. Cells forming long, straight filopodia (A-f for an example) were counted as the Cdc42 phenotype cells.

Fig. 5. F-actin binding/bundling activity of the IRSp53-IMD and MIM-IMD.

A. High-speed F-actin co-sedimentation assay. 1.2 µM GST-fused IRSp53 fragments (IRSp53-IMD and IRSp53-∆SH3), a MIM fragment (MIM-IMD), or GST alone (GST) were incubated with (+) or without (-) F-actin (5 µM as for G-actin) for 30 min on ice, then ultracentrifuged at 400,000 × g for 15 min. Equal aliquots of the supernatant (S) and the resuspended pellet (P) were analyzed by SDS-PAGE and Coomassie Blue staining.

B. Quantitation of F-actin binding activity of IMDs. Purified tag-free IMDs analysed by SDS-PAGE are shown in the left panel. 2 µM of tag-free IMDs and 0, 0.5, 1, 2, and 4 µM of F-actin were co-sedimented at 400,000 × g. The IMD bands stained with Coomassie Blue (right upper panel showing a representative data) were quantified by densitometry. Relatively long incubation time used in this experiment caused precipitation of 6 to 8% of the total IMDs without
F-actin. To show the net F-actin binding of the IMDs, the percentage of F-actin-bound IMD was calculated as the percentage recovered in the pellet subtracted by that recovered in the pellet without F-actin. Data shown in the lower panel are the mean ± SD of three independent experiments. S, supernatant; P, pellet.

C. Visualization of F-actin bundles induced by IMDs. F-actin (1.2 µM) was incubated with 1.2 µM GST-IRSp53-∆SH3 (a), 6 µM GST-IRSp53-IMD (b), 6 µM GST-MIM-IMD (c) or 1.2 µM GST alone (d) for 30 min on ice. After staining with Rhodamine-phalloidin, actin filaments and bundles were observed under a confocal microscope. Scale bar, 10 µm.

D. Quantitation of F-actin bundling activity of IMDs. 0, 0.5, 1, 2, 4, 8, and 16 µM of tag-free IMDs and 1 µM of F-actin were co-sedimented at 10,000 × g. 4 to 10% of total actin was recovered in the pellet in the absence of the IMDs. The percentages of actin in the bundles were quantified as described in Fig. 5B. Data shown in the lower panel are the mean ± SD of three independent experiments. S, supernatant; P, pellet.

Fig. 6. Electron micrographs of actin bundles induced by IRSp53.

A. Negatively stained F-actin bundles induced by GST-fused IRSp53-IMD for 30 min on ice. Scale bar, 1 µm.

B. Macroscopically visible tangled threads of actin bundles induced by GST-fused IRSp53-∆SH3 for 1 h on ice were packed by ultracentrifugation and processed for transmission electron microscopy. Transverse and longitudinal sections of actin bundles approximately 0.2 µm thick are shown. In a transversely sectioned area (inset, enlarged view of the rectangular area), many
actin filaments are aligned with a regular center-to-center distance of 11.2 ± 2.0 nm (mean ± SD, N = 854) as shown in Fig. 6C. Scale bar, 0.1 µm.

Fig. 7. Self-association of IMDs.

A. Analysis of cross-linked IMDs by SDS-PAGE. 1 µM of tag-free IMDs of IRSp53 and MIM were cross-linked with 4 mM EDC for 0, 15, 30 and 60 min. Chymotrypsinogen A was used as a non-dimer control. Equal amounts of protein were analyzed by SDS-PAGE followed by Coomassie Blue staining. An arrow and double arrows point to the monomer and the dimmer, respectively.

B. Co-immunoprecipitation assay for self-association of IMDs. GFP-tagged IMD, full length protein (WT) or ΔIMD was co-expressed with FLAG-tagged IMD, WT, ΔIMD or FLAG-vector alone (Vector) in 293T cells. GFP-tagged proteins were immunoprecipitated (IP) using anti-GFP antibody and analyzed by immunoblotting (IB) with anti-FLAG M2 antibody (upper panel). Expression levels of FLAG-tagged proteins (middle panel) and GFP-tagged proteins (lower panel) are shown by immunoblot analyses of total lysates.
Fig. 1

A

| Protein   | Region       | Alignment Score |
|-----------|--------------|-----------------|
| IRSp53    | 1            | MSLSRSBEMHRLETVN |...
| IRTKS     | 1            | MSRGPBENVNLRTETR |...
| FLJ22582  | 1            | MAPEDMQYRTSMATYK |...
| MIM       | 1            | MEAVIEKECSALGGL |...
| ABBA-1    | 1            | METAKECGALGGFQAT |...
| IRSp53    | 67           | QG-SKELGDVFQMAE |...
| IRTKS     | 66           | PV-STELECHVIEIS |...
| FLJ22582  | 66           | P-TSQILEGILVOS |...
| MIM       | 69           | RGCTREIGCSALTMC |...
| ABBA-1    | 68           | RGATRDIGCSALTMC |...
| IRSp53    | 135          | DKCQAELELKLRRKKS |...
| IRTKS     | 134          | EKSAQAELELKLRRKKS |...
| FLJ22582  | 134          | EKCMSLWMEKR      |...
| MIM       | 139          | KKSSDTL-KLQKKAKK |...
| ABBA-1    | 138          | KKSSDTL-KLQKKARK |...
| IRSp53    | 198          | VEKQCAVAKNSAAYH |...
| IRTKS     | 197          | VDKHCGFANHIHYHQL |...
| FLJ22582  | 188          | AEKHLLSNTFLQFF |...
| MIM       | 200          | SMLRPVIEEISMLGE |...
| ABBA-1    | 202          | TFLQPVVGELTMGE |...

B

| Protein   | IMD | CRIB | SH3 | WWB |
|-----------|-----|------|-----|-----|
| IRSp53    |     |      |     | 521 |
| IRTKS     |     |      |     | 511 |
| FLJ22582  |     |      |     | 518 |
| IRSp53-IMD| 250 |      |     |     |
| IRSp53-ΔSH3| 364 |      |     |     |
| IRSp53-GR | 521 |      |     |     |
| IRSp53-FF/AA| 251 |      |     |     |
| MIM       |     |      | 755 |
| ABBA-1    |     | 747  |     |     |
| MIM-IMD   |     |      | 250 |
| MIM-ΔIMD  |     |      | 251 |

C

IRSp53 sub-family (SH3 domain)

IRSp53
IRTKS
Gg IRSp53
Dr IRTKS
Hs IRSp53
Fl J22582
Gg IRTKS
Dr IRTKS

MIM/ABBA sub-family (WH2 domain)

Ce M04F3.5

Hs MIM
Gs MIM
Dr ABBA-2
Dr ABBA-1

Dm CG32082

Hs FLJ22582
Gs FLJ22582
Dr FLJ22582
**Fig. 3**

IRSp53-IMD  
IRSp53-IMD + Cdc42N17  
IRSp53-IMD + Rac1N17

(a) GFP  
(b) IRSp53-IMD  
(c) IRSp53-IMD + Cdc42N17  
(d) IRSp53-IMD + Rac1N17

**d**

| Treatment                  | % of filopodium forming cells |
|----------------------------|-------------------------------|
| GFP                        | 50                            |
| IRSp53-IMD                 | 60                            |
| IRSp53-IMD + Cdc42N17      | 70                            |
| IRSp53-IMD + Rac1N17       | 60                            |

Scale bar: 10 μm
Fig. 6

A

B

C

(distance between actin filaments in nm)

%
Fig. 7

A

![Image of a gel with bands indicating dimer and monomer formation](image)

B

| GFP-Tagged                | IRSp53       | MIM          |
|---------------------------|--------------|--------------|
| FLAG-IMD                  | IMD          | IMD          |
| FLAG-WT                   | WM           | WM           |
| FLAG-ΔIMD                 | AΔIMD        | AΔIMD        |
| Vector                    | -            | +            |

**IP:** α-GFP

**IB:** α-FLAG

**IB:** α-GFP
SUPPLEMENTAL DATA

Collection of nucleotide sequence data of putative members of the IRSp53/MIM-family proteins.

Cloning of human ABBA-1—Human ABBA-1 (F-actin-bundling protein with BAIAP2 homology in the N-terminal helical region, AB115770) was cloned from a human brain cDNA library (pDNR-LIB, Clontech) by PCR amplification using a pair of primers predicted from the human genomic sequence. The primers used were

5’AGAGCATGGAGACGGCGGAGAAGGAGTG (forward) and
5’CACCTGCTCGCACTGGGGCCTGAGAG (reverse) found in a hypothetical gene located at 16q22.1 (LocusID: 92154). The model mRNA of this gene, XM_208937, encodes a polypeptide similar to the C-terminal half of MIM. There are putative exons encoding amino acid sequences highly homologous to the N-terminal half of MIM in the upstream region. Expression of these putative exons is supported by expression sequence tags (ESTs) such as AL539489 and BU859607. The 2294 bp PCR product obtained has an open reading frame of 747 amino acids that shows 55% identity and 11% similarity to human MIM. The coding region of human ABBA-1 spans 22.2 kb and consists of 15 exons on the minus strand of a genomic contig NT_010498.13: join (19441131..19441197, 19436376..19436431, 19436201..19436273,
19435377..19435461, 19435192..19435283, 19435036..19435110, 19434720..19434728,
19433658..19433815, 19433198..19433305, 19429712..19429934, 19420872..19420946,
19420326..19420502, 19420004..19420169, 19419038..19419855). All the exons except one
keep AG/exon/GT boundaries and the exon 13 uses GC instead of GT. Northern blot analysis of
human ABBA-1 showed a single 6 kb band in various tissues with the highest expression level in
the brain.

**Cloning of Zebrafish IRSp53**—Zebrafish IRSp53 was PCR-amplified from Zebrafish adult brain
cDNAs using a pair of primers predicted from the zebrafish cDNAs and the genomic sequences
(An UniGene cluster Dr.16661, WGS 126238745 and a genomic contig BX323027). The forward
primer was 5’TCTCCTCCTCCTCCTCCTCTGTTTGTGTTTCTCG and the reverse primer was
5’AGCATGTCTGATAATGTGTATGAAAGGTG. The 1771 bp PCR product obtained has an
open reading frame of 534 amino acids that shows 63% identity and 17% similarity to human
IRSp53. The brain cDNAs were kindly gifted from Dr. M. Hibi (Riken Center for Developmental
Biology).

**Cloning of Zebrafish ABBA-1 and -2**—Zebrafish ABBA-1 and -2 were similarly cloned from the
brain cDNAs using two pairs of PCR primers predicted from the zebrafish database. We found
two N-terminal and two C-terminal sequences that are more closely related to ABBA-1 than
MIM. A forward primer (F1) 5’TTCTTCCAGCCGATTTTGTTCATTGAAAGGGAGCAG was
predicted from ESTs AL914911 and AL914912 and the other one (F2) 5’TGGGCTGGTAAAATGGATGCGGGAATGG was predicted from genomic sequences BX530017 and BX548167. Similar to human ABBA-1, the last exons of zebrafish putative ABBAs encode one third of the whole proteins, so that we could obtain two different relatively long 3’-sequences from AI959191 and overlapping WGSs and from BQ419430 and WGSs. The predicted 3’ primers were R1: 5’TGAAAAGCTTTATTTAGTGCGGGAATGG and R2: 5’AGGAGGCTCAGCAATCCGGGGTGCAGAG. Only the combinations of F1/R1 and F2/R2 successfully amplified PCR-products of 2240 bp and 2222 bp, respectively, and were tentatively named ABBA-1 and ABBA-2. The deduced aa sequence of zebrafish ABBA-1 contains 710 aa and shows 66% identity/9% similarity to human ABBA-1 and 51% identity/12% similarity to human MIM. ABBA-2 contains 737 aa that shows 73% identity/10% similarity to human ABBA-1 and 56% identity/11% similarity to human MIM.

**Database searches for putative IRSp53/MIM family proteins in vertebrates**—We found five candidates proteins in the human genome, IRSp53, IRTKS, FLJ22582, MIM and ABBA-1. To reveal the conservation of these proteins in vertebrates, we intensively searched Gallus gallus (chicken) sequences for the bird, Xenopus laevis (african clawed frog) and Silurana tropicalis (western clawed frog) sequences for the amphibia, and Danio reio (zebra fish) sequences for the
fish. Basically overlapping EST clones were used to make putative full-length mRNA sequences. Mismatches and possible sequence errors were corrected by overlapping ESTs and genomic sequences, mainly whole genome shotgun (WGS) sequences. We have found that the exon compositions of these genes are almost completely conserved and more than five WGSs could be obtained for most of the exons. GenBank accession numbers of representative clones used are sited below.

**IRSp53 sub-family**

Human IRSp53 isoform 1. NP_059344.

Mouse IRSp53. NP_570932.

Bovine IRSp53. UniGene Bt.5610; ESTs, BF654505, BI847652, AV602077, BF604061, BF040137, AV612697, BI541686, BE721861.

Chicken IRSp53. BM440642, BG712911, CD218196, BM491233, BI847652.

*Silurana tropicalis* IRSp53. UniGene Str.6299; the N-terminal fragments, AL892953, AL847406. The C-terminal fragment, AL864224.

Human IRTKS. NP_061330.

Mouse IRTKS. NP_080109.

Bovine IRTKS. the C-terminal fragment, UniGene Bt.11916; BF044090, BF046149, BF045001, CB437083.
Chicken IRTKS. BU473774, BU129548, BX262287, BU293404.

Zebra fish IRTKS. AAH50238, UniGene Dr.11420.

Human FLJ22582. NP_079321.

Mouse FLJ22582. AAH48937.

Chicken FLJ22582. CD728910, BU307330, CD733604, BU372164, BU305484.

Xenopus FLJ22582. The C-terminal fragment, UniGene Xl.19687, CA972469, CA973430, BQ383816, BQ383815.

Zebra fish FLJ22582. The N-terminal fragment, AL916895, CA974969, CD595810, BI980054.

**MIM sub-family**

Human MIM. NP_055566.

Mouse MIM. NP_659049.

Bovine MIM. The C-terminal fragment, AW437626, BF043708, BF076606, BE476284.

Chicken MIM. BU446345, BU359736, BU359378, BU131818, BI066132, BM426009, BU257939.

Xenopus MIM. The N-terminal fragment, BU913061, BU913016, BQ726165. The C-terminal fragment, BU916773, BI444844.

Mouse ABBA-1. Partial sequence was found in the locus 244654 on chromosome 8, whereas mouse MIM is encoded by the locus 211401 on chromosome 15. The model sequence,
XM_146515, was extended and edited by a number of ESTs such as CB525721, CD355768, CD348498 and AK052172. There may be splice isoforms. The sequence shown is the closest form to our human ABBA-1 clone. There is a rat ortholog, XM_226503. chicken ABBA-1. The N-terminal fragment, BU372800, BU142478, BU396053, BU372986.
Supplemental Fig. 1. Alignment of IRSp53 proteins.

| Species | IRSp53 | Amino Acid Sequence |
|---------|--------|---------------------|
| Hs      | IRSp53 | MSLRSEEMHRLTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |
| Mm      | IRSp53 | MSLRSEEMHRLTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |
| Bt      | IRSp53 | MSLRSEEMHRLTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |
| Gg      | IRSp53 | MSAEDEVRHRTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |
| St      | IRSp53-N| MSLRSEDEVRHRTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |
| Dr      | IRSp53 | MSLRSEDEVRHRTENYVTMEQPNLSLRNFIAKMNGYKALAGTVAYAAKGYFDALVKMGELASESQGSK |

The IMD is boxed. The half-CRB motif, the SH3 domain and the WW binding motif are shown in a black box, shaded areas and a shaded box, respectively. Asterisks show the identity and dots show the similarity. Species: Hs, human; Mm, mouse; Bt, bovine; Gg, chicken; St, Silurana tropicalis; Dr, zebra fish. For detail, see Collection of nucleotide sequence data of putative members of the IRSp53/MIM family proteins.
The IMD is boxed. The SH3 domain is shaded and the WW binding motif is in a shaded box. Asterisks show the identity and dots show the similarity. Species: Hs, human; Mm, mouse; Bt, bovine; Gg, chicken; St, *Silurana tropicalis*; Dr, zebrafish. For detail, see Collection of nucleotide sequence data of putative members of the IRSp53/MIM family proteins.
Supplemental Fig. 3. Alignment of FLJ22582 proteins.

The IMD is boxed. The SH3 domain is shaded. Asterisks show the identity and dots show the similarity. FLJ22582 proteins show lower levels of sequence conservation than other members in the family.

Species: Hs, human; Mm, mouse; Gg, chicken; Xl, Xenopus laevis; Dr, zebra fish. For detail, see Collection of nucleotide sequence data of putative members of the IRSp53/MIM family proteins.
Supplemental Fig. 4. Alignment of MIM sub-family proteins.

| Hs MIM | 71 | GTRIEGSLTRMCRRHRSIEAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Mm MIM | 71 | GTRIEGSLTRMCRRHRSIEAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Gg MIM | 71 | GTRIEGSLTRMCRRHRSIEAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Xl MIM | 71 | ATKIEGSLTRMCRRHRIIESRLHRHFTALMEMQIEWKEKVRANQLDKHAEYKQAQE1KKK |
| Hs ABBA-1 | 70 | ATPQIGSLTRMCRRHRSITAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Mm ABBA-1 | 70 | ATRQIGSLTRMCRRHRSITAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Gg ABBA-1 | 60 | ATRQIGSLTRMCRRHRSITAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Dr ABBA-1 | 60 | ATRQIGSLTRMCRRHRSITAKRQRFSPALIDCINPQLQMEWKEKVRANQLDKHAEYKQAQE1KKK |
| Dr ABBA-2 | 71 | ATKIEGSLTRMCRRHRIIESRLHRHFTALMEMQIEWKEKVRANQLDKHAEYKQAQE1KKK |

| Hs MIM | 141 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Mm MIM | 141 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Gg MIM | 141 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Xl MIM | 141 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Hs ABBA-1 | 140 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Mm ABBA-1 | 140 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Gg ABBA-1 | 140 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Dr ABBA-1 | 140 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |
| Dr ABBA-2 | 141 | SSDTLKQKKAKK---GRGDIQPQLDSALQDVNDYLLLEETEKLAVRKILIEERGRFCTFILRPVIE |

| Hs MIM | 208 | EISMLGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Mm MIM | 208 | EISMLGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Gg MIM | 208 | EISMLGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Xl MIM | 208 | EISMLGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Hs ABBA-1 | 210 | GELTMGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Mm ABBA-1 | 210 | GELTMGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Gg ABBA-1 | 210 | GELTMGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Dr ABBA-1 | 210 | GELTMGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |
| Dr ABBA-2 | 210 | GELTMGIEQLSLTHQIESEDLKLTPMKPLSSQVILDDKLSYSSYQTPSPSSPTMSSKRSVCSSSL |

| Hs MIM | 278 | NSVNSSDSSRSS---GSHSHPSSHYRYSRNLSAQAPV-RLSVSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Mm MIM | 278 | NSVNSSDSSRSS---GSHSHPSSHYRYSRNLSAQAPV-RLSVSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Gg MIM | 278 | NSVNSSDSSRSS---GSHSHPSSHYRYSRNLSAQAPV-RLSVSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Xl MIM | 278 | NSVNSSDSSRSS---GSHSHPSSHYRYSRNLSAQAPV-RLSVSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Hs ABBA-1 | 280 | SSSSASAKGQAFPGAQTQPSSTCTYRSLQAPATT-TLSVSSSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Mm ABBA-1 | 280 | SSSSASAKGQAFPGAQTQPSSTCTYRSLQAPATT-TLSVSSSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Gg ABBA-1 | 280 | SSSSASAKGQAFPGAQTQPSSTCTYRSLQAPATT-TLSVSSSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Dr ABBA-1 | 280 | SSSSASAKGQAFPGAQTQPSSTCTYRSLQAPATT-TLSVSSSSHDSGFQSDQA-FQGKSSPSPMPPEA |
| Dr ABBA-2 | 278 | SAHSSASRSS---GQGQTHSSPSHCYRSSLQAPVTTAHLRSVSSSSHDSGFQSDQA-FQGKSSPSPMPPEA |

| Hs MIM | 343 | PNQLSNGFHYSSLYSSESVHGPTGAQFHPCLRPRLPRTRSVTVSHPDHYAHYITGCMF---PSSQIPS |
| Mm MIM | 343 | ANQLSNGFHYSSLYSSESVHGPTGAQFHPCLRPRLPRTRSVTVSHPDHYAHYITGCMF---PSSQIPS |
| Gg MIM | 343 | PNQLSNGFHYSSLYSSESVHGPTGAQFHPCLRPRLPRTRSVTVSHPDHYAHYITGCMF---PSSQIPS |
| Hs ABBA-1 | 347 | TSQKSSASSASASRSSETQVCVS---ECSSPTACS |
| Mm ABBA-1 | 347 | TSQKSSASSASASRSSETQVCVS---ECSSPTACS |
| Dr ABBA-1 | 347 | TSQKSSASSASASRSSETQVCVS---ECSSPTACS |
| Dr ABBA-2 | 346 | TSQKSSASSASASRSSETQVCVS---ECSSPTACS |
Supplemental Fig. 4. Alignment of MIM sub-family proteins (page 2).

| Protein | Sequence |
|---------|----------|
| Hs MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Mm MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Gg MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Xl MIM-C| RLDQSGIPCI- |
| Hs MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Mm MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Gg MIM  | GDDQAEAD-QQEQDKSSTIFPRSNDSIQYRMRMFKARPASTAGLPTTGLT-PMVTV-PGVATI |
| Hs ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |
| Mm ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |
| Gg ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |
| Xl ABBA-1 | 11 RLDQSGIPCI- |
| Hs ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |
| Mm ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |
| Gg ABBA-1 | 411 SPAASDLAVMTRLGLSLRQCKSLSQGYSSTTTTPCSEDTPISQVS-PMVTV-PGVATI |

The IMD is boxed. The WH2 domain is underlined. Asterisks show the identity and dots show the similarity. MIM and ABBA are highly homologous, but characteristic amino acid differences are noted even in highly conserved regions (shaded residues in MIM sequences). The N-terminus of chicken ABBA-1 could not be retrieved. Two zebrafish cDNAs show the ABBA characters. Species: Hs, human; Mm, mouse; Gg, chicken; Xl, Xenopus laevis; Dr, zebrafish. For detail, see Collection of nucleotide sequence data of putative members of the IRSp53/MIM family proteins.
Supplemental Fig. 5. Amino acid sequence alignment of the IRSp53/MIM homology domains.

Species used: Hs, human; Gg, chicken; Dr, zebra fish; Dm, *D. melanogaster*; Ce, *C. elegans*. Red lines are the consensus helices predicted and black dots and lines represent conserved amino acid residues. Residues color codes: acidic, red; basic, blue; alcoholic, light blue; aromatic, brown; polar, green; P, black; C, yellow; non-polar, plain. The insertion of four amino acids in Dr IRSp53 (see Supplemental Fig. 1) is not shown.
Supplemental Fig. 6. Amino acid sequence homology in the C-terminus of the IRSp53/MIM family proteins.

Species: Hs, human; Gg, chicken; Xl, *Xenopus laevis*; Dr, Zebra fish
Supplemental Table 1. The IMD containing proteins in the human genome.

| Gene symbol | LocusID | Chromosomal position | Model protein | Helical contents in IMD | pI | Homology between proteins | Homology between IMDs |
|-------------|---------|----------------------|---------------|-------------------------|----|--------------------------|-----------------------|
| IRSp53      | BAIAP2  | 10458                | 17q25         | NP_059344               | 84%| 9.0                      | -                     | 22/18%                |
| IRTKS       | LOC55971| 55971                | 7q22.1        | NP_061330               | 82%| 8.8                      | b                     | 46/28%                |
| FLJ22582    | FLJ22582| 80115                | 22q13.1       | NP_079321               | 84%| 9.2                      | b                     | 28/29%                |
| MIM         | MTSS1   | 9788                 | 8p22          | NP_055566               | 99%| 8.5                      | -                     | 22/18%                |
| ABBA-1      | LOC92154| 92154                | 16q22.1       | XP_208937a              | 99%| 8.8                      | c                     | 17/21%                |

*a* C-terminal half; *b* Comparison to IRSp53, identity/similarity; *c* Comparison to MIM.
A novel actin-bundling/filopodium-forming domain conserved in insulin receptor tyrosine kinase substrate p53 and missing in metastasis protein
Akiko Yamagishi, Michitaka Masuda, Takashi Ohki, Hirofumi Onishi and Naoki Mochizuki

J. Biol. Chem. published online January 29, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M309408200

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

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

Supplemental material:
http://www.jbc.org/content/suppl/2004/02/03/M309408200.DC1