A Diverse Family of Proteins Containing Tumor Necrosis Factor Receptor-associated Factor Domains*

We have identified three new tumor necrosis factor-receptor associated factor (TRAF) domain-containing proteins in humans using bioinformatics approaches, including: MUL, the product of the causative gene in Muckle-Wells syndrome; USP7 (Hausp), an ubiquitin protease; and SPOP, a POZ domain-containing protein. Unlike classical TRAF family proteins involved in TNF family receptor (TNFR) signaling, the TRAF domains (TDs) of MUL, USP7, and SPOP are located near the NH2 termini or central region of these proteins, rather than carboxyl end. MUL and USP7 are capable of binding in vitro via their TDs to all of the previously identified TRAF family proteins (TRAF1, TRAF2, TRAF3, TRAF4, TRAF5, and TRAF6), whereas the TD of SPOP interacts weakly with TRAF1 and TRAF6 only. The TD of MUL also interacted with itself, whereas the TDs of USP7 and SPOP did not self-associate. Analysis of various MUL and USP7 mutants by transient transfection assays indicated that the TDs of these proteins are necessary and sufficient for suppressing NF-κB induction by TRAF2 and TRAF6 as well as certain TRAF-binding TNF family receptors. In contrast, the TD of SPOP did not inhibit NF-κB induction. Immunofluorescence confocal microscopy indicated that MUL localizes to cytosolic bodies, with targeting to these structures mediated by a RBCC tripartite domain within the MUL protein. USP7 localized predominantly to the nucleus, in a TD-dependent manner. Data base searches revealed multiple proteins containing TDs homologous to those found in MUL, USP7, and SPOP throughout eukaryotes, including yeast, protists, plants, invertebrates, and mammals, suggesting that this branch of the TD family arose from an ancient gene. We propose the moniker "TRAFs (TD-encompassing factors) for this large family of proteins.

Tumor necrosis factor (TNF) receptor-associated factors (TRAFs) constitute a family of adapter proteins first identified for their binding to TNF family receptors (TNFRs). TRAF family proteins regulate several of the functions of the TNFR superfamily, apparently by linking the cytosolic domain of those receptors to downstream protein kinases or ubiquitin ligases (1–3). Six members of the TRAF family (TRAF1 through TRAF6) have been identified to date in humans and mice (1, 2).

All mammalian TRAFs and two recently characterized Drosophila TRAFs (4, 5) share a distinctive region near their COOH terminus denominated the “TRAF domain” (TD). The TD represents a novel protein fold of about 180 amino acids that is responsible for the interaction of the TRAFs with the TNF receptors and other adaptors and kinases (1). The x-ray crystallographic structures of the TDs of TRAF2 and TRAF3 show that they form a seven-stranded anti-parallel β-sandwich structure, with each TRAF domain monomer containing a surface crevice responsible for binding peptideyl motifs found in the cytosolic domains of the TNF family receptors to which they bind (6–10). Similarities and differences in the peptideyl specificities of individual TRAFs account for their selective associations with particular TNFR family members and other TRAF-binding proteins, yielding specificity, diversity, redundancy, and competition among TRAFs with respect to ligand-inducible recruitment to various TNF family receptor complexes (10–17). Through interactions with other adapter proteins, TRAFs also indirectly associate with members of the interleukin-1 receptor/Toll family (18, 19).

Several TRAFs can also bind a variety of protein kinases, including the NF-κB-inducing kinases IRAKs, NIK, RIP1, and RIP2/CARDIACK and the c-Jun NH2 kinase pathway activators MEKK1, Ask1, Missiphan (Msn), and the Germinal Center kinase-related kinase (4, 5, 20–27). Thus, TRAF proteins physically and functionally connect TNFRs and interleukin-1R/Toll receptors to intracellular protein kinases, thereby linking these receptors to downstream signaling pathways. However, not all TRAFs are capable of interacting with downstream kinases, and some may function therefore as antagonists of TNFR and interleukin 1-receptor/Toll signaling (28, 29).

Furthermore, some TDs self-associate, forming trimeric structures stabilized both by complementary in the eight-stranded β-sandwich fold that constitute the COOH-terminal region of TDs and by a NH2-terminal coiled-coil region that stabilizes TD trimerization (6–9). Heterotypic interactions among TDs have been described (16, 30), suggesting the possibility of mixed trimers that theoretically may account for the antagonism displayed among certain TRAF family members.

Meprins are a class of dimeric extracellular metalloproteinases of the astacin family (31) that also contain putative TDs of unknown function (32). Since the region in the meprins that...
TEFs: A Diverse Family of TRAF-containing Proteins

shares homology with the TD is extracellular, it is unlikely that it plays any role in regulating TNF family receptors or intracellular TRAF family proteins. The presence of TD-like sequences in mpeins, however, suggests additional and broader roles for the TDs in cell physiology (32).

In this report, we described a new branch of the TRAF family, which includes in humans the proteins MUL, USP7, and SPOP. The MUL gene has recently been shown to be mutated in patients with Mulibrey Nanism, an autosomal recessive disorder that affects several tissues of mesoderm origin (33). USP7 is a ubiquitin-specific protease that has been reported to bind the Vmw110 protein of herpes simplex virus (HSV1) (34, 35). SPOP was previously identified as an autoantigen in a patient with scleroderma pigmentosum (36). All three of these new TRAF proteins have a different topological organization compared with the classical TRAFs implicated in TNF-R and Toll signaling, with the TDs located internally (MUL) or in the NH2-terminal part of the molecule (USP7 and SPOP), instead of the customary COOH-terminal localization seen in the other previously identified human TRAFs.

As supportive evidence that these proteins do indeed contain TDs, we demonstrate that these new TRAFs are able to interact with classical TRAF family proteins in vitro and to modulate NF-kB induction by them, at least when overexpressed in cells. However, the physiological role of these novel TD-containing proteins may be directed to purposes other than TNFR- and interleukin 1-receptor/Toll-mediated signal transduction, as discussed herein. Finally, using bioinformatics approaches, we identified several putative TD-containing proteins in diverse eukaryotic species, including yeast, protozoa, slime molds, nematodes, arthropods, and plants, suggesting an early evolutionary origin of the TD and implying roles beyond cytokine receptor signal transduction. We propose the moniker TEFs (TD-encapping factors) for this large family of proteins.

MATERIALS AND METHODS

Computer Methods—The PSI-BLAST program (NCBI, National Institute of Health) was used for searches of the public data bases using as a query the predicted amino acid sequence of the TD of hTRAF2. Multiple sequence alignments were performed using the standard PM250 parameters from ClustalX and ClustalW. Relations among the different members of the TD family were calculated using the phylogenetic algorithm implemented in ClustalX. Models of the TDs of MUL, USP7, and SPOP were constructed by threading on the structure of the TD of TRAF2 (6, 7), using FFAS and MODELER (37, 38).

Multiple sequence alignments were performed using the standard PM250 parameters from ClustalX and ClustalW. Relations among the different members of the TD family were calculated using the phylogenetic algorithm implemented in ClustalX. Models of the TDs of MUL, USP7, and SPOP were constructed by threading on the structure of the TD of TRAF2 (6, 7), using FFAS and MODELER (37, 38).

Production and Purification of Recombinant TRAF Domains—For GST fusion protein production, pGEX-plasmids were transformed into competent XL-1 blue bacteria cells and grown in LB medium. When bacteria reached an OD600 of 1.0, GST-protein production was induced with 1 mM isopropyl-1-thio-b-galactopyranoside for 4 h at 25 °C. Cells were then recovered and resuspended in phosphate-buffered saline containing 1 mM diithiothreitol, 1 mM phenylmethylsulfonyl fluoride, and 100 μg/ml lysozyme and lysed by sonication. The GST fusion proteins were purified from bacterial lysates by affinity chromatography using glutathione-Sepharose (Amersham Pharmacia Biotech), essentially as described (9, 16). The resins were then washed with phosphate-buffered saline containing 1 mM diithiothreitol until the OD600 of the wash reached <0.01.

Protein Binding Experiments—In vitro GST-protein binding assays were performed as described (4, 16). Briefly, [35S]methionine-labeled TRAF proteins were synthesized by coupled in vitro transcription-translation (Promega Inc.). Equal amounts of each labeled protein (2–6 μg/ml) were added to rabbit reticulocyte lysates with luciferase as described (9, 16). The resins were then washed with phosphate-buffered saline containing 1 mM diithiothreitol until the OD600 of the wash reached <0.01.

Reporter Gene Assays—For NF-kB reporter gene assays, 293T cells were transfected using a calcium phosphate method and a total of 12 μg of plasmid DNA (including 0.5 μg of pUC13-4xNFkB-luc plasmid and 1 μg of pCMV-β-galactosidase plasmid) at ~50% confluence in 6-well plates in duplicate. After 36 h, cells were lysed with 0.5 ml of Promega lysis buffer. The luciferase activity of 10 μl of each cell lysate was determined using the Luciferase assay system from Promega, following the manufacturer’s protocol, and measured using a luminometer (EG&G Berthold). Luciferase activity was normalized relative to β-galactosidase activity (mean ± S.E.).

Immunofluorescence Confoocal Microscopy—COS7 cells were transfected with LipofectAMINE Plus (Life Technologies, Inc.) and a total of 3 μg of DNA. At 24 h after transfection, cells were plated in complete tissue culture medium onto poly-lysinated cover glasses and allowed to settle for 24 h at 37 °C, 5% CO2. Cells were fixed with 1:1 (v/v) methanol/acetone. After preblocking in 10 mM Hepes, pH 7.4, 150 mM NaCl, 3% bovine serum albumin, 1% goat serum, cells were incubated with anti-Myc mAb (Santa Cruz Biotechnology) followed by secondary fluorocesin isothiocyanate-labeled rabbit anti-mouse IgG (Dako) and incu-
bated with 1 μg/ml propidium iodide. Confocal microscopy was performed using a two-photon system (Bio-Rad).

RESULTS

Identification of Three Human TEFs—BLAST searches of the public databases, including the non-redundant database of protein sequences maintained at NCBI and the HGTS database, were performed using the sequence of the TD of human TRAF2 as a query. Three human cDNAs displaying significant homology with the TRAF2 TD were thus identified: MUL (3e−19), USP7 (e−17), and SPOP (3e−24). The putative TDs of MUL, USP7, and SPOP share ~31–38% (mean 33.6%) amino acid sequence identity with each other, and ~9–23% (mean 16.7%) sequence identity with the TDs of human (hu) TRAF1–TRAF6 (Fig. 1). In contrast, the TDs of TRAF1–TRAF6 share ~31–69% (mean 47.4%) amino acid sequence identity with each other (Fig. 1B).

To further interrogate the possibility that MUL, USP7, and SPOP contain TDs, threading approaches were used, taking

---

**Fig. 1.** Comparison of sequences and predicted structures of TRAF domain proteins. A, an amino acid sequence alignment (ClustalW, standard PAM250 parameters) is presented for the TRAF domains of human USP7, SPOP, MUL, and TRAF1–6. Black and gray boxes indicate identical and similar (conserved) residues, respectively. B, comparative analysis of the percentage of identical amino acid residues present in various TDs. C, computer-generated models of TRAF domains are presented for human MUL, USP7, SPOP, and TRAF2, using the coordinates of the x-ray crystal structure of human TRAF2 as a template for threading. The coiled-coil of TRAF2 bears no sequence similarity to the coiled-coil of MUL and therefore could not be modeled using the TRAF2 coordinates, but it was added to the figure to illustrate the similarities in topology between TRAF2 and MUL. D, a schematic representation of the three human TEF-family proteins, MUL, USP7, and SPOP, is presented, making comparisons with TRAF1 and TRAF2. Symbols are as indicated.
advantage of the published structure of the TD of huTRAF2 (6, 7). Indeed, the sequences corresponding to the TDs of MUL, USP7, and SPOP readily adapted to the TD fold (Fig. 1C), without evidence of internal inconsistencies. All 3 of these proteins were predicted to contain the eight-stranded anti-parallel β-sandwich structure of the COOH-terminal portion of the TD (so-called “TRAF-C” domain). In addition, MUL contains a predicted long α-helical domain upstream of the TRAF-C domain, analogous to TRAF2, and thus also displays predicted structural similarity to the classical TRAFs in the so-called “TRAF-N” region.

Examination of the predicted complete ORFs of MUL, USP7, and SPOP reveals that the TDs are located either internally in the molecule (MUL) or near the NH2 terminus (USP7 and SPOP) (Fig. 1D). Therefore, these proteins have a different topological organization from the previously described members of the TRAF family, which all have their TRAF domains located near the COOH terminus.

Similar to the classical TRAFs, the MUL protein also contains a RING finger domain (aa 15–55) near its NH2 terminus (Fig. 1D). This is followed by a ZF-Box domain (aa 90–132) and then a predicted α-helical coiled-coil region (132–177), forming a tripartite motif, denominated the RBCC domain for RING, B-box, and coiled-coil. RBCC domains have been previously recognized in a variety of proteins (43). After the RBCC domain, a second coiled-coil region (aa 195–231) is found in MUL. Two leucine zipper domains may also be present in this region of the protein (aa 197–218 and 222–245) (not shown), which resides upstream of the TD. After the TD (aa 273–403), another predicted coiled-coil region (aa 427–446) is found, followed by two segments which are rich in acidic amino acid residues (aa 452–577 and 868–964). This COOH-terminal region of MUL distal to the TD has weak amino acid sequence similarity to murine histone deacetylase 1 and Haloarcula marismortui ribosomal protein L12, as determined by PSI-BLAST searches (44) (not shown). Interestingly, MUL contains two putative nuclear localization signals (PAVEKRR at aa 847 and KRRK at aa 851).

The predicted USP7 protein is 1102 aa in length (135 kDa), and contains two domains sharing homology with ubiquitin-specific proteases (aa 215–233 and 448–518), which reside downstream of the predicted TD (aa 58–196) (Fig. 1D). Indeed, this region of USP7 has been confirmed by biochemical assays to possess ubiquitin-specific protease activity (34).

SPOP is a 374-amino acid protein of unknown function which contains a POZ domain (aa 190–289) downstream of the predicted TD (aa 33–164). This protein has been previously localized to nuclear bodies in a speckled pattern, thus prompting the acronym SPOP for Speckle-type POZ-domain protein (36).

The TDs of MUL, USP7, and SPOP Can Bind Classical TRAFs—TDs from the classical TRAF family proteins are capable of interacting with themselves and sometimes each other. We therefore explored whether the predicted TDs of MUL, USP7, and SPOP could bind in vitro to human TRAF1, TRAF2, TRAF3, TRAF4, TRAF5, TRAF6, and the TRAF-binding protein I-TRAF (TANK) (45, 46). For these experiments, the TDs of MUL, USP7, and SPOP were expressed as GST fusion proteins in bacteria and purified by glutathione-Sepharose affinity chromatography, then tested for interactions with 35S-labeled in vitro translated TRAF1–6, I-TRAF, and the TDs of MUL, USP7, and SPOP (Fig. 2). Consistent with the structural predictions, the TDs of MUL and USP7 interacted in vitro with all of the classical TRAFs (TRAF1–6). The TD of SPOP also interacted with certain TRAFs, specifically TRAF1 and TRAF6. The TD of MUL additionally interacted with itself and with I-TRAF, but did not bind the TDs of USP7 or SPOP (Fig. 2). The TDs of USP7 and SPOP, in contrast, failed to self-associate, unlike classical TRAFs which commonly form trimers. Of note, the region of MUL expressed for these experiments contained only the predicted 8-strand anti-parallel β-sheet region, but lacked the upstream predicted α-helical segments that forms coiled-coils and stabilizes trimerization of classical TRAFs. Binding experiments using GST-control proteins as well as GST-MUL, USP7, and SPOP (TDs) in combination with a variety of irrelevant control proteins confirmed the specificity of these results (not shown).

Since TRAF family proteins are known to bind the cytosolic domains of certain members of the TNFR family, we also explored the possibility that MUL, USP7, or SPOP similarly might interact with these receptors. Accordingly, in vitro protein interaction assays were performed in which the TDs of MUL, USP7, and SPOP (as well as TRAF2, which was employed as a positive control) were tested for interactions with GST fusion proteins containing the cytosolic domains of several TNF family receptors. The TDs of MUL, USP7, and SPOP were either produced as 35S-labeled proteins by in vitro translation or generated by expression as epitope-tagged proteins in HEK293T cells using transient transfection methods. Among these proteins, only the TD of MUL displayed interactions with TNF family receptors in vitro (Fig. 3 and data not shown). These data raise the possibility that the TD of MUL contains sufficient structural similarity with classical TRAF family proteins to recognize the motifs within the cytosolic domains of some TNF family receptors, at least in vitro. Again, binding experiments using GST control proteins as well as GST-Fas, TNFR2, CD40, LTαR, and NGFRp75 in combination with a variety of irrelevant control proteins confirmed the specificity of these results (not shown).

**MUL and USP7 Modulate NF-κB Induction by TRAFs—**

Because transient overexpression of some TRAFs induces NF-κB activity, we tested the effects of MUL, USP7, and SPOP on activation of a NF-κB reporter gene plasmid by transient transfection assays in HEK293 cells. Overexpression of MUL, USP7, or SPOP failed to induce NF-κB (not shown). Instead,
MUL and USP7 inhibited NF-κB induction caused by transient overexpression of TRAF2 and/or TRAF6. As shown in Fig. 4, for example, co-transfection of plasmids encoding full-length MUL or USP7 with TRAF2 at a 2:1 ratio (TEF:TRAF) suppressed NF-κB activity by approximately half, relative to cells transfected with plasmids encoding TRAF2 alone. MUL also inhibited NF-κB induction by TRAF6, although USP7 was less effective against this TRAF family member.

Moreover, the TDs of MUL and USP7 were sufficient for NF-κB inhibition. In fact, expression of a TD-containing region of USP7 in the absence of its ubiquitin-protease domain was more inhibitory than full-length USP7, suppressing NF-κB induction by both TRAF2 and TRAF6 (Fig. 4). In contrast, the TD of SPOP failed to suppress NF-κB induction caused by overexpression of either TRAF2 or TRAF6. Likewise, expressing deletion mutants of MUL and USP7 lacking their TDs failed to suppress NF-κB induction. Immunoblot analysis indicated that MUL and USP7 (as well as their TDs) did not interfere with TRAF2 or TRAF6 protein production (not shown). We conclude therefore that the TDs of MUL and USP7 are necessary and sufficient for suppression of NF-κB induction in this assay where overexpression of TRAFs is used as a mechanism for triggering an NF-κB response.

**Subcellular Localization of MUL and USP7—TRAFs are typically found in the cytosol, reportedly localizing to cytosolic structures of undetermined origin (47–49). We therefore explored the intracellular locations of MUL and USP7, and several deletion mutants of these proteins, using immunofluorescence confocal microscopy and COS-7 cells transiently transfected with plasmids encoding Myc epitope-tagged MUL or USP7. As shown in Fig. 5, MUL is associated with cytosolic bodies, which appear as dots throughout the cytosol (panel A). In addition to COS-7 cells, this same pattern of immunolocalization was also seen in a variety of other cell lines, when transfected with plasmids producing Myc-tagged or GFP-tagged MUL, including, HT1080, HT20, HeLa, and 293T. Two-color analysis using antibodies specific for proteins of mitochondria, Golgi, lysosomes, or megasomes failed to suggest an organelle to which MUL targets (not shown). Moreover, MUL did not co-localize with TRAF2, TRAF6, or the TNF family receptors CD40, Fas, and DR5.2

To determine the regions within the MUL protein responsible for this pattern of subcellular targeting, a variety of MUL deletion mutants were expressed as Myc-tagged proteins by transient transfection. Mutants lacking the RING domain (panel B), containing only the RBCC domain (panel C), lacking the polyacidic region (panel D), or lacking the TRAF domain (panel E) displayed a similar subcellular location compared with full-length MUL. Thus the RBCC domain is sufficient for targeting to punctate cytosolic structures. Furthermore, within the RBCC, the RING domain is expendable for proper targeting. In contrast, diffuse cellular immunofluorescence encompassing both the nucleus and cytosol was obtained for cells expressing Myc-tagged fragments of MUL containing only the TD (panel F), the TD with adjacent coiled-coil region (panel G), or TD with the adjacent COOH-terminal domain containing the polyacidic segments and candidate nuclear localization signal sequence (NLS) (panel H). A fragment of MUL containing

\[ \text{TD} \]

\[ \text{RING FINGER} \]

\[ \text{COILED COIL} \]

\[ \text{POLYACIDIC DOMAIN} \]

\[ \text{NLS} \]

\[ \text{TRAF DOMAIN} \]

\[ \text{PROTEASE DOMAIN} \]

\[ \text{Zn BOX} \]

\[ \text{polyacidic regions and candidate nuclear localization signal sequence (NLS)} \]

\[ \text{(panel H)} \]

\[ \text{A fragment of MUL containing} \]

---

\[ J. M. Zapata and J. C. Reed, unpublished observations. \]
plasmids containing the MUL(TEF3) mutants depicted at the right-side of the figure: (A) full-length MUL; (B) MUL (aa 63–964); (C) MUL (aa 1–270); (D) MUL (aa 1–414); (E) MUL 254–450; (F) MUL (aa 265–414); (G) MUL (aa 206–414); (H) MUL (aa 265–964); and (I) MUL (aa 412–964). At 48 h after transfection, cells were fixed in methanol-acetone and stained with anti-Myc mAb, followed by anti-mouse IgG-fluorescein isothiocyanate and propidium iodide. Cells were imaged by confocal microscopy. Cells stained with secondary antibody and cells transfected with empty pcDNA3-myc plasmid served as negative controls (not shown). 100-μm bars are shown in the Fig.

Fig. 5. Subcellular localization of MUL and MUL deletion mutants. COS7 cells were transfected in 6-well dishes with 3 μg of pcDNA3-myc plasmids containing the MUL(TEF3) mutants depicted at the right-side of the figure: (A) full-length MUL; (B) MUL (aa 63–964); (C) MUL (aa 1–270); (D) MUL (aa 1–414); (E) MUL 254–450; (F) MUL (aa 265–414); (G) MUL (aa 206–414); (H) MUL (aa 265–964); and (I) MUL (aa 412–964). At 48 h after transfection, cells were fixed in methanol-acetone and stained with anti-Myc mAb, followed by anti-mouse IgG-fluorescein isothiocyanate and propidium iodide. Cells were imaged by confocal microscopy. Cells stained with secondary antibody and cells transfected with empty pcDNA3-myc plasmid served as negative controls (not shown). 100-μm bars are shown in the Fig.

only the COOH-terminal domain with polyacidic segment and NLS was localized primarily to nuclei, although some fainter cytosolic immunofluorescence was also seen (Fig. 5f).

Similar experiments were performed for USP7. The full-length USP7 protein was located predominantly in the nucleus, although cytosolic immunofluorescence was also observed (Fig. 6A). A deletion mutant of USP7 containing essentially only the TD (aa 1–212) displayed a similar pattern of immunostaining, except the cytosolic component was more pronounced than seen with the full-length USP7 protein. In contrast, a mutant of USP7 lacking the TRAF domain (aa 202–1102) (panel B) was completely excluded from the nucleus and instead was located diffusely throughout the cytosol. These results therefore suggest that the TD of USP7 is necessary and sufficient for target of this protein to nuclei.

TEF Proteins Have an Ancient Origin and Can Be Classified into Three Groups—PSI-BLAST searches of the public data bases were seeded with the sequences of the TDs of MUL, USP7, and SPOP, as well as the TDs of TRAF1–6, and run to saturation in search of additional TEFs. Only predicted polypeptides yielding e values < 0.001 after five iterations were considered positive. These studies revealed the existence of multiple independent cDNAs encoding potential TEF family proteins in nearly all eukaryotic lineages, from yeast to humans. Currently, almost 260 homologous proteins can be identified with high confidence as members of the extended TEF family. The species containing candidate TEFs include yeast (Schizosaccharomyces pombe and Saccharomyces cerevisiae), protozoa (Trypanosoma brucei), Dicytostelium (D. discoideum), nematodes (Caenorhabditis elegans), insects (Drosophila melanogaster), plants (including monocots (Oryza sativa (rice), Shorgum bicolor), and dicots (Arabidopsis thaliana and Medicago truncatula)), amphibians (Xenopus laevis), and mammals (human, mouse, and rat). It is interesting to note that despite the complete sequencing of several prokaryotic genomes, including examples of both bacteria and archaea, no TEFs were found in these organisms.

An alignment of the amino acid sequences of some of the TDs of these TEF family proteins is presented in Fig. 7. To make the alignment more readable, it was simplified by representing groups of closely homologous proteins with one representative example (a complete multiple sequence alignment and accompanying phylogenetic tree are available from our WEB server). Eight blocks of high homology are found within the aligned TD sequences, corresponding to the β-strands of the TDs of TRAF2 and TRAF3 (6–10). The sequence alignment also demonstrates the remarkably high percentage of identity of the TDs of human SPOP and its counterparts in Drosophila (CG9924) (96% identity) and C. elegans (YNV5) (94% identity). Although only the TD sequences are presented in Fig. 7, comparison of the complete ORFs of the SPOP proteins of humans, flies, and worms also reveals striking sequence similarity (Drosophila SPOP has 79% identity (89% homology) and C. elegans SPOP has 63% identity (79% homology) to human SPOP). This level of conservation among such distant species suggests an important role for SPOP in animal cell physiology.

Next, the phylogenetic program was used to produce a dendrogram in which different groups of TEFs with higher homologies could be identified (Fig. 8). It is important to note that because of the use of orthologs, paralogs, and several proteins from the same species in the tree, accurate reflection of the actual evolutionary distances among the different members of the TEF family is not possible. Relations among the different members of the TEF family were calculated using the standard parameters of the phylogenetic programs implemented in ClustalX. The use of the Fitch-Margoliash method (Phylip) produced similar results (not shown). The TEF proteins represented in the tree can be subdivided in three groups. Group I contains the previously known TRAF proteins, including human TRAF1 through TRAF6, and the recently identified TRAFs from Drosophila and C. elegans. This group I branch also includes the Meprins, a family of extracellular metallopro-
III typically have a domain organization differing from the ancient, since they are present in yeast and protozoa, in addition to plants or unicellular organisms, with the closest found in C. elegans, Arabidopsis (At), Oryza (Or), and Shorgum (Sb)). Group II includes USP7, several USP7-related proteins that contain TDs in combination with apparent ubiquitin-specific protease domains (Drosophila (CG1490), C. elegans (3878045 and 3877391), Arabidopsis (6671947), and yeast, including S. pombe and S. cerevisiae (6014652, UBPB-SCHPO and UBPF_yeast), and additional TEFs from Drosophila, C. elegans, plants, Dicyostelium, Trypanosoma, and C. elegans. We presume that the Group III TEs are the most ancient, since they are present in yeast and protozoa, in addition to higher organisms. Proteins belonging to Groups II and III typically have a domain organization differing from the Group I proteins, with the TRAF domain located near the NH₂ terminus, followed by a diversity of COOH-terminal domains.

**DISCUSSION**

*Identification of Three TEFs in Humans—* Using bioinformatics methods, we have identified three proteins in humans, MUL, USP7, and SPOP, that contain a variant of the TD. The evidence that this conserved domain recognized in the MUL, USP7, and SPOP proteins represents a variant of the TD includes (a) sequence alignments; (b) structure prediction (threading); (c) ability of variant TDs to bind the TDs of classical TRAF family members; and (d) modulation of the function of classical TRAFs in terms of NF-κB induction. In the case of MUL, this similarity to classical TRAFs also extends to additional protein interactions, including an ability to bind I-TRAF (TANK) and some TNF family receptors. Unlike the previously described members of the TRAF family, the TDs in MUL, USP7, and SPOP are located near the NH₂ terminus rather than COOH terminus. Expression studies indicate that the mRNAs encoding MUL, USP7, and SPOP are widely present in adult human tissues (Refs. 33 and 36, and data not shown), suggesting the functions of these TEF-family proteins are likely to be applicable to multiple cell types.

USP7 (TEF1)—USP7 (HAUSP) contains a variant TD followed by two ubiquitin-specific protease (USP) domains. Among the 3 human TEF family genes described here, it was the first to be identified and thus we have proposed TEF1 as an alternative designation for this protein. The protease domain-containing region of this USP7 has been expressed in bacteria and confirmed by biochemical assays to be capable of cleaving polyubiquitin chains (34). The USP7 (TEF1) gene maps to chromosome 16p13.3 (50), representing a region commonly involved in translocations, deletions, and other cytogenetic abnormalities in cancers such as endymonomas, basal cell carcinomas, and acute myeloid leukemias (51–53). Genes within the 16p13 region have also been implicated in certain hereditary syndromes, including Robinstein-Taybi syndrome, tuberous sclerosis complex, adult polycystic kidney disease, and α-thalassemia/familial mental retardation syndrome (54–59). However, it is unknown whether USP7 (TEF1) is directly involved in cancers or genetic diseases. Homologues of USP7 (TEF1) which contain both a TD and USP domain are present in diverse organisms, with the closest found in Drosophila (CG1490), which shares >60% amino acid sequence identity with its human counterpart. This strong conservation of amino acid sequence homology suggests an evolutionarily conserved function for this protein.

USP7 was originally identified based on its ability to associate with the herpes simplex virus-type 1 immediate-early protein Vmw110 (34). Vmw110 localizes to nuclear structures where PML is found, the so-called PML oncogenic domains (PODs). The function of PODs is unclear, but they have been implicated in interferon responses and pathogenesis of various human diseases, including acute promyelocytic leukemia and viral infections (reviewed in Ref. 60). Increasing evidence suggests a role for PODs as specialized sites of transcriptional regulation, perhaps in assembling multiprotein transcriptional complexes or controlling targeted degradation of nuclear proteins (61–64). The Vmw110 protein is a potent activator of gene expression and is required for efficient virus reactivation from latency (35). Of potential relevance to its interactions with a USP, the Vmw110 protein reportedly activates proteosome-dependent degradation of several substrates, and it has been suggested that this targeted protein degradation is required to re-activate the lytic cycle of the virus (65, 66). Expression of Vmw110 in cells initially causes an apparent increase in the levels of Vmw110, followed by re-activation of the lytic cycle of the virus (65, 66). Expression of Vmw110 in cells initially causes an apparent increase in the Vmw110 protein, followed by re-activation of the lytic cycle of the virus (65, 66). Expression of Vmw110 in cells initially causes an apparent increase in the Vmw110 protein, followed by re-activation of the lytic cycle of the virus (65, 66).
between Vmw110 and USP7 also inhibit the activation of gene expression and virus replication mediated by Vmw110, suggesting a causal role for USP7 as a co-factor in these virus-mediated processes (35).

Since USP7(TEF1) is normally located primarily in the nucleus, it is unlikely that this protein participates in early signal transduction events triggered by TNF family receptors and TRAFs. It is conceivable, however, that USP7(TEF1) could...
regulate or could be regulated by TRAFs that translocate to the nucleus under certain circumstances, such as TRAF4 which is reported to reside in the nuclei of some tumors (67). Interestingly, our deletional analysis suggests that the TD is responsible for the nuclear targeting of USP7 (TEF1), raising the possibility that this domain binds other nuclear proteins. Overexpression of a truncation mutant of USP7 (TEF1) containing essentially only the TD may have saturated nuclear-binding sites, explaining the spillover of this protein fragment into the cytosol and its potent suppression of TRAF-induced NF-κB. Although recruitment of a USP to TRAFs might be expected to interfere with NF-κB induction by preventing IκB degradation resulting from polyubiquitination (68), the USP domains of USP7 (TEF1) are not required for TRAF antagonism.

**SPOP (TEF2)**—The SPOP (TEF2) protein was previously identified as an autoantigen in a patient with scleroderma pigmentosum (36). The TD of SPOP (aa 33–164) is located near the NH2 terminus of the molecule, and is followed by a POZ domain (aa 190–289). POZ domains have been identified in several transcriptional repressors (reviewed in Ref. 69). These domains bind components of the SMRT-NCoR repressor complex, which has histone deacetylase activity (70, 71). SPOP (TEF2) and several of the POZ family proteins localize to discrete nuclear structures, but whether these are the same as PODs remains controversial. Deletional analysis of the SPOP (TEF2) protein has provided evidence that both the TD and POZ domain are required for targeting to nuclear speckles (36). Of note, the TDs of SPOP and USP7 failed to interact in our in vitro protein binding assays, suggesting that these proteins do not directly associate, despite their targeting to nuclear subcompartments. Although the TD of SPOP displayed at least weak interactions with TRAF1 and TRAF6 in vitro, it did not demonstrate an antagonistic effect on TRAF-mediate induction of NF-κB. The physiological roles of SPOP (TEF2) thus remain to be defined. The striking amino acid sequence similarity of the human, *Drosophila*, and *Caenorhabditis* homologues of SPOP (TEF2) suggests a conserved function within the animal kingdom. The human SPOP gene is located in chromosome 17, and is flanked by NGFR (TNFR-16) and distal-less homeobox 4 genes.

**MUL (TEF3)**—The MUL (TEF3) protein contains a RBCC domain, TD, and polyacidic domains. Unlike the other TEFs identified thus far in humans, MUL (TEF3) is a cytosolic protein which is localized in a distinctive punctuate pattern. Furthermore, the size of these punctate structures increased, with increasing levels of MUL achieved by transfection. Based on immunofluorescence analysis of deletion mutants of MUL (TEF3), we determined that the RBCC domain is necessary and sufficient for targeting to cytosolic structures. The identity of these foci of MUL (TEF3) accumulation, however, is unclear. Using two-color immunofluorescence techniques, we have excluded mitochondria, lysosomes, Golgi, and megasomes as likely candidates for the cytosolic bodies with which MUL (TEF3) associates. Interestingly, the subcellular localization of MUL (TEF3) is reminiscent of some other members of the RBCC family, such as BERP and Rfp (72, 73). Deletional analysis indicates that the B-box of Rfp is required for its targeting to cytosolic granule-like structures (73), raising the possibility that elements within the RBCC tripartite domain account for targeting of MUL, Rfp, BERP, and certain other members of the RBCC family to these cytosolic structures. Also, the BERP protein binds myosin V and α-actinin, and has been speculated to regulate organelle transport (72, 74). It remains to be determined, however, whether MUL (TEF3) co-localizes with BERP or Rfp in cells. Moreover, the location within cells of some RBCC family members such as Rfp (which can function as a repressor of the HIV LTR) and Xnf (which is involved in dorsal-ventral patterning in *Xenopus*) can change from cytosolic to nuclear in concert with differences in cell context or protein modification by phosphorylation (73, 75). Thus, the location of MUL (TEF3) may be subject to regulation, although we consistently observed a punctate cytosolic pattern of immunofluorescence in 5 of 5 tumor cell lines examined here.

Based on its subcellular location, MUL (TEF3) conceivably could participate in physiological regulation of TRAF family proteins. Of note, previous attempts to localize TRAF family proteins have revealed association with punctate cytosolic structures prior to activation of TNF family receptors, followed by translocation of TRAFs to the plasma membrane after ligand addition (47–49). However, two-color immunofluorescence studies indicated that: (a) TRAF2 and TRAF6 do not co-localize with MUL (TEF3) and (b) the subcellular targeting of MUL did not correlate with suppression of TRAF function, inasmuch as a truncation mutant of MUL containing essentially only the TD was equally effective as the full-length protein, despite its failure to target to cytosolic subregions. It might also be noted that the TD-only mutant of MUL (TEF3) was also found in the Triton X-100 soluble fraction, while...
full-length MUL(TEF3) was predominantly in the Triton X-100 insoluble fraction. Thus, the association of MUL(TEF3) with the Triton X-100 insoluble fraction precluded attempts to assess association of this protein with classical TRAFs by co-immunoprecipitation assays. At this point, therefore, no compelling evidence exists to imply a physiological role of MUL(TEF3) in regulating signal transduction by TRAFs, although additional experimentation will be required to evaluate this possibility further.

While this paper was in preparation, the MUL(TEF3) gene on chromosome 17q22–23 was reported to be mutated in patients with Mulibrey Nanism, an autosomal recessive disorder that affects several tissues of mesodermal origin (33). Mulibrey Nanism is characterized by severe growth failure of prenatal onset, constrictive pericardium with consequent hepatomegaly, hypoplasia of several endocrine glands with consequent hormonald deficiency, triangular face with hydrocephaloid skull, and susceptibility to develop Wilm’s tumors. A substantial portion of patients are suspected to be lost by early abortion and others by infantile death. Thus, the MUL(TEF3) protein plays important roles in human development and possibly tumor suppression, but the biochemical mechanism of the protein remains enigmatic.

Four independent frameshift mutations in the MUL(TEF3) gene have been identified in families with Mulibrey Nanism, all of which are predicted to produce truncated versions of the MUL(TEF3) protein. Although it remains to be determined whether these mutant MUL proteins are stable when expressed in cells, the shortest of the truncated proteins identified thus far retains only the RING, B-box, and a portion of the coiled-coil domain, thus having an incomplete RBCC tripartite domain. Two mutants retain the RBCC domain and TD, while another lacks only the last 227 amino acids and thus is missing the second of the two polyacidic domains (33). With respect to the truncation mutants of MUL(TEF3) associated with Mulibrey Nanism, our deletion analysis suggests that at least 3 of the 4 mutant proteins would still target to cytosolic structures, since the RBCC domain was sufficient for localization to punctate cytosolic structure in immunofluorescence experiments. Thus, it is unlikely that improper subcellular targeting uniformly accounts for the dysfunction of such mutant MUL(TEF3) proteins. Future explorations of the protein interaction partners of the MUL(TEF3) protein may provide insights into the specific biochemical defect that accounts for the diverse developmental abnormalities seen in patients harboring mutations of their MUL(TEF3) genes.

TEFs Represent an Ancient Group of Proteins Found in Diverte Unicellular and Multicellular Eukaryotes—The discovery of multiple proteins containing candidate TDs in organisms as diverse as yeast, protozoa, plants, nematodes, flies, amphibians, and mammals suggests a very ancient origin for this protein family. The biochemical mechanism of the MUL(TEF3) protein remains enigmatic.

REFERENCES
1. Arch, R. H., Gedrich, R. W., and Thompson, C. B. (1998) Genes Dev. 12, 2921–2930
2. Wallach, D., Varfolomeev, E. E., Malinina, N. L., Golovtsev, Y. V., Kovalenko, A. V., and Boldin, M. P. (1999) Annu. Rev. Immunol. 17, 331–367
3. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., Yoo, J., Slatter, D., Pickart, C., and Chen, Z. J. (2000) Cell 102, 351–361
4. Zapata, J., Matsuzawa, S., Godzik, A., Lee, E., Wasserman, S., and Reed, J. (2000) J. Biol. Chem. 275, 12102–12107
5. Liu, H., Su, Y.-C., Becker, E., Treisman, J., and Skolnik, E. (1999) Curr. Biol. 9, 101–104
6. McWhirter, S., Pullen, S., Holton, J., Crute, J., Kehry, M., and Albert, T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 8408–8413
7. Park, Y., Burkitt, V., Villa, A., Tong, L., and Wu, H. (1999) Nature 398, 533–538
8. Ye, H., Park, Y., Kriekman, M., Kieff, E., and Wu, H. (1999) Mol Cell 4, 321–330
9. Ni, C.-Z., Welsh, K., Lee, E., Chiao, C.-K., Wu, H., Reed, J. C., and Ely, K. R. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 16395–16399
10. Park, Y. C., Ye, H., Hata, C., Segal, D. R., Rich, R. L., Hou, H.-C., Myszka, D. G., and Wu, H. (2000) Cell 101, 777–787
11. Gedrich, R. W., Gibillini, M. C., Ducket, C. S., Van Dongen, J. L., and Thompson, C. B. (1998) J. Biol. Chem. 273, 12852–12858
12. Boucher, L.-M., Munegre, E. M., Lu, Y., Thukral, S., and Tak, M. W. (1997) Biochem. Biophys. Res. Commun. 233, 592–600
13. Pullen, S., Miller, H., Everdeen, D., Dang, T., Crute, J., and Kehry, M. (1998) Biochemistry 37, 11830–11845
14. Pullen, S., Labadia, M., Inagami, R., McWhirter, S., Everdeen, D., Alber, T., Crute, J., and Kehry, M. (1999) Biochemistry 38, 10168–10177
15. Pullen, S. S., Dang, T. T., Crute, J. I., and Kehry, M. R. (1999) J. Biol. Chem. 274, 14246–14254
16. Lee, E., Welsh, K., Matsuzawa, S., Zapata, J. M., Kitada, S., Mitchell, R., Ely, K. R., and Reed, J. C. (1999) J. Biol. Chem. 274, 22414–22274
17. Yamamoto, H., Kishimoto, T., and Minamoto, S. (1998) J. Immunol. 161, 4753–4759
18. Muzio, M., Natoli, G., Saccani, S., Levervo, M., and Mantovani, A. (1998) J. Exp. Med. 187, 2097–2108
19. Cao, Z., Xiang, J., Takeuchi, M., Kurama, T., and Goddell, D. V. (1996) Nature 383, 443–446
20. Hsu, H., Huang, J., Shu, H., Baichwal, V., and Goddell, D. V. (1996) Immunity 5, 387–396
21. Song, H. Y., Regnier, C. H., Kirschning, C. J., Goddell, V. C., and Rotter, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 9792–9796
22. Sasa, S., Leonardti, A., Kyrion, J., Siebenlist, U., and Kehrl, J. H. (1999) J. Immunol. 163, 3279–3285
23. Thome, M., Hofmann, K., Burns, K., Martinson, F., Bodmer, J., Mattman, C., and Teppoch, J. (1998) J. Biol. Chem. 273, 885–888
24. McCarthy, J., Ni, J., and Dixit, V. (1998) J. Biol. Chem. 273, 16968–16975
25. Hoeßlich, K. K., Yeh, W. C., Yao, Z., Mak, T. W., and Woodgett, J. R. (1999) Oncogene 18, 5814–5824
26. Ishihara, H., Saitsu, M., Mochida, Y., Takeda, K., Nakano, H., Roth, M., Miyazono, K., and Ichijo, H. (1999) Mol. Cell 2, 389–395
27. Baud, V., Liu, Z.-G., Bennett, B., Suzuki, N., Xia, Y., and Karin, M. (1999) Genes Dev. 13, 1297–1308
28. Rothe, M., Sarma, V., Dixit, V. M., and Goddell, D. V. (1995) Science 269, 3276–32770
29. Ye, X., Mehlen, P., Rabizadeh, S., VanArsdalen, T., Zhang, H., Shim, H., Wang, J., Lee, E., Zapata, J., Hauser, C., Reed, J., and Bredeisen, D. (1999) J. Biol. Chem. 274, 30202–30208
30. Rothe, M., Wang, S. C., Henzel, W. J., and Goddell, D. V. (1994) Cell 78, 681–692
