Interaction between TFF1, a Gastric Tumor Suppressor Trefoil Protein, and TFIZ1, a Brichos Domain-Containing Protein with Homology to SP-C†

Bruce R. Westley,‡,§ S. Michael Griffin,|| and Felicity E. B. May*,‡,§

Department of Pathology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne NE1 4LP, U.K., Northern Institute for Cancer Research, School of Clinical and Laboratory Sciences, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K., and Northern Oesophago Gastric Cancer Unit, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne NE1 4LP, U.K.

Received December 24, 2004; Revised Manuscript Received February 22, 2005

ABSTRACT: TFF1 is a gastric tumor suppressor that protects gastric epithelial cells from damage but can promote invasive properties of tumor cells. Antibodies were raised against correctly folded TFF1 protein. These showed that the 6.67 kDa secreted trefoil protein is present as an ∼25 kDa complex in normal human gastric mucosa. The TFF1 complex was immunopurified from human gastric mucosa and shown to comprise two proteins joined by a disulfide bond. Both were identified by amino-terminal sequencing and MALDI TOF mass spectrometry. The TFF1 protein partner is a previously unknown protein that we have called TFIZ1 for *t*refoil *f*actor *i*nteractions(*z*) 1. TFIZ1 is expressed and secreted in normal gastric mucosa. TFIZ1 mRNA was cloned from gastric mucosa and sequenced. TFIZ1 is an 18.31 kDa protein and contains an ∼100 amino acid brichos domain and homology with smart00019.10, SF_P. This is the first demonstration that a member of the trefoil factor family of proteins is bound covalently to a brichos domain-containing protein. The apparent molecular mass of the TFF1:TFIZ1 heterodimer is remarkably close to the theoretical molecular mass of 24.98 kDa. In conclusion, the heterodimer comprises one molecule each of TFF1 and TFIZ1, and the disulfide bond between TFF1 and TFIZ1 is the most important factor stabilizing the heterodimer.

Gastric cancer has a propensity to metastasize and kills 640000 people worldwide each year. TFF1[1] is a 6.67 kDa secreted protein originally identified as an estrogen-regulated mRNA in breast cancer cells (*1*, *2*). The main site of expression of TFF1 in normal tissues is the stomach (*3*, *4*). TFF1 is a member of the trefoil factor family of proteins (*5*−*7*). Trefoil proteins are synthesized by mucin secreting epithelia and are important in mucosal protection and repair. TFF1 interacts with the soluble human gastric mucin MUC5AC (*8*) and has been proposed as a receptor for the class I carcinogen *Helicobacter pylori* which colonizes the adherent mucus layer (*9*). Studies with transgenic mice, and administration of recombinant protein, show that TFF1 protects from experimentally induced gastrointestinal damage (*10*, *11*). TFF1 is thought to promote restitution (*11*), during which epithelial cells adjacent to a damaged area flatten, extend lamellipodia, and migrate to re-form an intact epithelial barrier.

Work with TFF1-null mice shows that TFF1 is a tumor suppressor. The gastric mucosa is thickened at 3 weeks (*12*) and by 5 months exhibits severe hyperplasia and dysplasia and absence of mucin. All TFF1-null mice then develop antropyloric adenomas and 30% develop adenocarcinomas (*12*). TFF1 expression is lost frequently in human gastric cancer (*13*, *14*), and missense mutations have been identified in 16% of carcinomas (*15*). It has been reported that *TFF1* gene inactivation can occur by loss of heterozygosity (*16*), and hypermethylation of the *TFF1* promoter has been reported in 19% of gastric cancers (*17*).

In apparent contradiction, TFF1 is frequently overexpressed or expressed ectopically by carcinoma cells (*6*). It stimulates migration and invasion of tumor cells (*18*, *19*) and is thought to be involved in tumor dissemination. TFF1 has antiapoptotic activity (*20*) and promotes angiogenesis (*21*). It is not known why TFF1 expression is beneficial in normal gastric cells and prevents the development of gastric tumors but promotes carcinogenesis and invasion in other cell types.

Trefoil proteins contain a 42−43 amino acid trefoil domain that has conserved features including six cysteine residues that form three intramolecular disulfide bonds (*5*−*7*). TFF1 contains one trefoil domain and is 60 amino acids long (Figure 1A). It has an extra cysteine residue near the carboxy terminus that participates in intermolecular interactions (Figure 1A,B) (*22*, *23*). In contrast to the defined loop structure of the trefoil domain, the carboxy-terminal portion of the TFF1 monomer is unstructured (*24*). In TFF1

† This work was supported by the Wellcome Trust, Cancer Research U.K., and the Newcastle University Hospitals' Special Trustees.

* Address correspondence to this author at The Medical School. Tel: 00 44 191 246 4417. Fax: 00 44 191 246 4301. E-mail: F.E.B.May@ncl.ac.uk.

‡ Department of Pathology, University of Newcastle upon Tyne.

§ Northern Institute for Cancer Research, The Medical School.

|| Northern Oesophago Gastric Cancer Unit, Royal Victoria Infirmary.

[1] Abbreviations: TFF, trefoil factor family; TFIZ1, trefoil factor interactions 1; SDS, sodium dodecyl sulfate; BSA, bovine serum albumin; PVDF, poly(vinylidene difluoride); KLH, keyhole-limpet hemocyanin; SP-C, pulmonary surfactant-associated protein C precursor; SF_P, pulmonary surfactant protein.

10.1021/bi047287n CCC: $30.25
Published on Web 05/12/2005

homodimers (*23*), two monomer units are connected by a flexible linker comprising the carboxy-terminal regions with an intermolecular disulfide bond between the Cys58 residues (*25*). The TFF1 dimer is more potent than the monomer in vitro and in animal models (*11*, *18*).

The majority of human gastric TFF1 is present in a complex or heterodimer of ~25 kDa (*26*). This has profound implications for TFF1 function. The protein partner in the heterodimer may confer specificity or affinity to the binding of TFF1 to a receptor or other proteins or affect the ability of TFF1 to stimulate restitution. It is likely to influence the tumor suppressor activity of TFF1. We report the purification and identification of the TFF1 protein partner as a previously unknown secreted protein. TFF1 is present in the heterodimer as a monomer unit that is linked to its protein partner by a disulfide bond.

EXPERIMENTAL PROCEDURES

Cell Culture. AGS cells were maintained in Ham's F12/Dubecco's modified Eagle's medium supplemented with 10% fetal calf serum. Kato-III cells were maintained in RPMI medium supplemented with 20% fetal calf serum. HGT-1 cells, HT29 cells, and NS-1 cells were maintained in Dubecco's modified Eagle's medium supplemented with 10% fetal calf serum.

Preparation of TFF1 Antibodies. TFF1 peptides of between 10 and 30 amino acids corresponding to different regions of TFF1 that were predicted to be immunogenic were synthesized. The peptides were conjugated to KLH, and mice were immunized using established protocols. Recombinant TFF1 was produced as described previously (*23*). Correctly folded TFF1 was purified, and the TFF1 homodimer was allowed to form and then purified. Mice were immunized with TFF1 dimer. Hybridoma cell lines were produced by fusion of mouse spleen cells to NS-1 cells. Screening for production of monoclonal antibodies was by ELISA.

Protein Gel Electrophoresis and Western Transfer Analysis. Acrylamide gels were prepared and proteins electrophoresed and transferred to PVDF membranes or stained with Coomassie blue essentially as described previously (*23*). Proteins were boiled for 5 min in the absence or presence of β-mercaptoethanol prior to electrophoresis. Filters were incubated with TFF1 monoclonal antibodies diluted in 5% milk, followed by incubation with alkaline phosphatase conjugated secondary antibodies, and the immunoreaction was developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as described previously (*27*).

Immunoprecipitation and Purification of the TFF1 Heterodimer. Cytosols were preabsorbed with protein G−Sepharose and incubated with TFF1 monoclonal antibodies for 4 h and subsequently with colloidal magnetic protein G beads in 300 mM NaCl, 100 mM Tris-HCl, pH 8.0, and 0.2% Nonidet-P40 at 4 °C overnight. The magnetic beads were retained in a MACS column (Miltenyi Biotec) in a magnetic field and washed with buffer, and the immunoprecipitated proteins were eluted with 62.5 mM Tris-HCl, pH 6.8, 12.5 mM EDTA, 10% (v/v) glycerol, 0.005% Bromophenol blue, and 1% SDS at 90 °C.

For preparative scale immunoprecipitations, hybridoma cells were cultured in protein-free medium, and then TFF1 monoclonal antibodies secreted into the medium were purified on protein A−Sepharose and conjugated to protein G−Sepharose using disuccinimidyl suberate. Incubations were as above, and the antibody-conjugated protein G beads were washed in buffer and the immunoprecipitated proteins eluted.

Protein Sequencing. TFF1 immunoprecipitate was electrophoresed on a polyacrylamide gel and transferred to PVDF membrane, and the protein bands were stained with 0.2% Coomassie blue R and destained in 40% methanol:10% acetic acid. The filter was dried, the appropriate bands were excised, and the amino-terminal sequences were obtained by gas-phase sequencing.

Matrix-Assisted Laser Desorption Time-of-Flight (MALDI TOF) Mass Spectrometry. TFF1 immunoprecipitates were electrophoresed and the proteins stained with Coomassie blue. Protein-containing gel slices were destained in 50% acetonitrile and 100 mM ammonium bicarbonate, dried, rehydrated in NH_4HCO_3 and 10% acetonitrile containing 0.5 μg of Trypsin Gold (Promega), and incubated overnight at 37 °C. The tryptic digests were then recovered from the gel slices using 50% acetonitrile and 5% trifluoroacetic acid. In some experiments, proteins were reduced and alkylated using dithiothereitol and iodoacetic acid prior to electrophoresis to avoid interaction of free cysteine residues with acrylamide during electrophoresis. Mass spectra were obtained from untreated samples or following further purification with Zip Tip pipet tips (Millipore) using an Applied Biosystems MALDI TOF system.

RT PCR. Total RNA was extracted from the precipitate obtained after homogenization of tissue or cells in LiCl and urea as described (*2*). RNA was reprecipitated twice with 150 mM KCl and 75% ethanol and dissolved in water at 1 mg/mL. For reverse transcription, 0.5 μg of RNA was incubated with 0.5 μg of random primers in 5.9 μL for 10 min at 70 °C, cooled on ice, and then incubated in 50 mM Tris-HCl, pH 8.3, 79 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 500 μM dNTPs, and 0.5 unit of RNAsin (Amersham) with 10 units of Moloney murine leukemia virus reverse transciptase (Amersham) in 10 μL for 1 h at 37−42 °C. PCR reactions contained 0.1 μL of the reverse transcription reaction in a final volume of 12.5 μL with Red Hot Polymerase or Amplitaq Gold (Applied Biosystems) as recommended by the manufacturer. The primers 5′-ATAAGGGTGCTGTTTCGAC-3′ and 5′-GTCAGAGCAGTCAATCTGTG-3′ amplified the 248 bp sequence of TFF1 cDNA from exon 2 to exon 3 (*28*). The primers 5′-TGCAGGATCATGCTCTTCTACC-3′ and 5′-CTCCAGAGGGTTGTACTTGACC-3′ amplified the 211 bp sequence of TFIZ1 cDNA from exon 3 to exon 5. The primers 5′-TCAATCCCTTGATGCACTGG-3′ and 5′-TCCATAGTACACTGGTCG-3′ amplified the 236 bp sequence of TFIZ2 (CA11/AMP-18) cDNA from exon 4 to exon 6 (*29*). PCR products were analyzed on 3% agarose gels.

DNA Sequencing. RT PCR was performed as above with oligonucleotides 5′-TTACTCCAGCACCTTCCTCTCC-3′ and 5′-ATAAATACAGCATTTATATTTTCTGAC-3′ to amplify the 718 bp sequence of TFIZ1 cDNA from exon 1 to exon 6 and with oligonucleotides 5′-ATAACACCTAGTTTGAGTCAACC-3′ and 5′-TGAATTCAATGCTAAATGATTTTATTG-3′ to amplify the 820 bp sequence of TFIZ2 (CA11/AMP-18) cDNA from exon 1 to exon 6 (*30*). The PCR products were extracted with chloroform and cloned

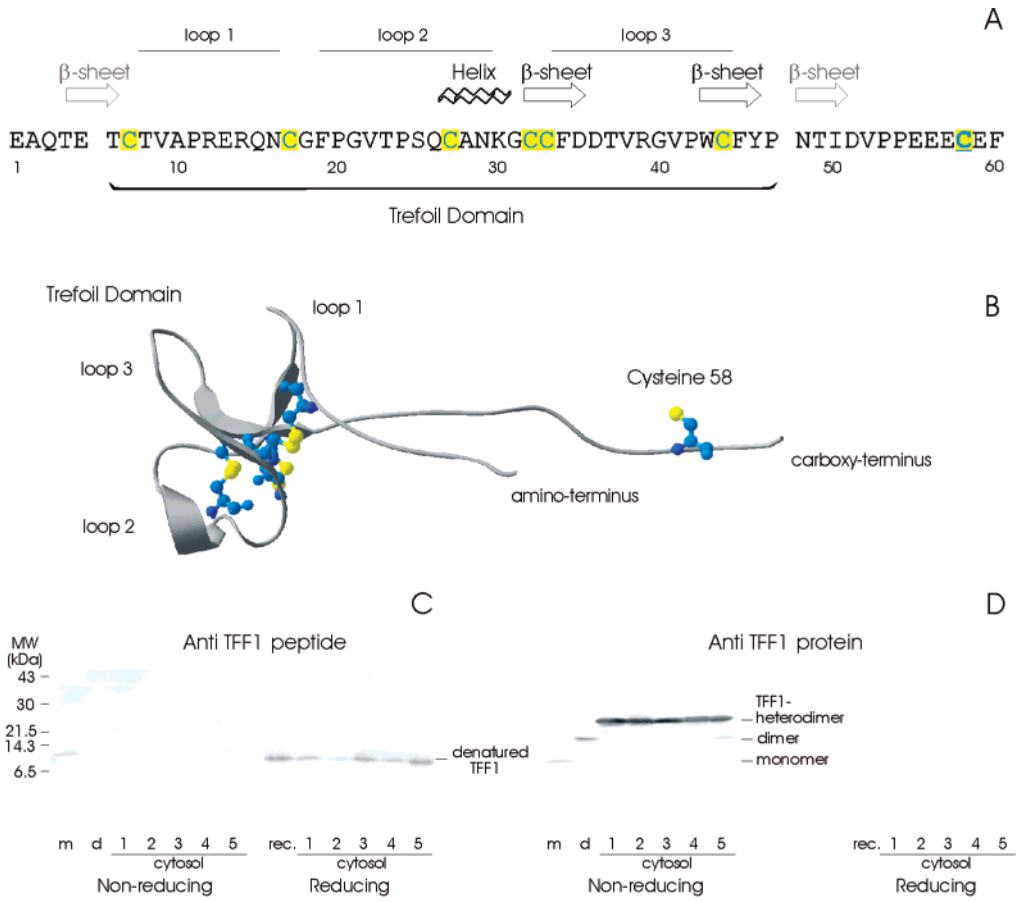


FIGURE 1: Production of monoclonal antibodies against correctly folded TFF1. (A) Amino acid sequence for human TFF1 showing the extent of the trefoil domain (residues 6−47). The conserved cysteine residues in the trefoil domain and the extra trefoil domain cysteine residue of TFF1 are shown in blue and are highlighted in yellow. In the TFF1 homodimer, an intermolecular disulfide bond connects the Cys58 residues (underlined). The positions of the secondary structural elements (β-sheets, helix, and loops) present in TFF1 are indicated above the sequence. The β-sheet formed between residues 3−6 and 48−50 is shown in gray because it was only present in a proportion of the NMR structures (*24*). (B) Location and orientation of the conserved cysteine residues that form intramolecular disulfide bonds in the TFF1 trefoil domain and of cysteine 58 in the carboxy terminus that forms intermolecular disulfide bonds. The cysteine residues are shown in blue apart from the sulfur atoms, which are shown in yellow. (C, D) Immunodetection of recombinant TFF1 and of TFF1 in human gastric mucosa cytosol. Samples were electrophoresed on polyacrylamide gels without or after prior incubation with β-mercaptoethanol. Separated proteins were transferred to PVDF membrane and incubated with a monoclonal antibody raised against a TFF1 peptide (C) or against recombinant TFF1 protein (D). Recombinant TFF1 monomer (m), recombinant TFF1 dimer (d), or denatured recombinant TFF1 (rec.) was included in the gels as indicated. The positions of the molecular mass markers are shown on the left and of the different forms of TFF1 on the right of the appropriate panels.

into the vector pCR2.1 by TOPO cloning. Cloned cDNA was sequenced by Big Dye termination cycle sequencing (Applied Biosystems).

RESULTS

Purification of the TFF1 Heterodimer

Production of Monoclonal Antibodies against Native TFF1. Monoclonal antibodies were raised against TFF1 peptides and subsequently against recombinant protein that had been used to solve the structure of TFF1 (*23*−*25*). The ability of the antibodies to recognize linear TFF1 and different forms of correctly folded TFF1 was tested. Representative results obtained are shown in Figure 1. After electrophoresis under nonreducing conditions the antipeptide antibody reacts weakly with the TFF1 monomer but not with the TFF1 dimer or TFF1 from gastric mucosa. The antibody does react strongly with recombinant TFF1 and gastric TFF1 after pretreatment of the proteins with β-mercaptoethanol (Figure 1C). The inability of the anti-TFF1 peptide antibodies to recognize folded TFF1 renders them ineffective for immunopurification of the TFF1 heterodimer.

Antibodies raised against correctly folded TFF1 protein react with the recombinant TFF1 monomer and dimer and with the TFF1 heterodimer and dimer in gastric cytosol after the proteins have been electrophoresed under nonreducing conditions but do not react with TFF1 if it is reduced prior to electrophoresis (Figure 1D). The TFF1 heterodimer is the predominant form of TFF1 detected in all five samples of gastric mucosa. Antibodies that react with native forms of TFF1 were used to immunopurify the TFF1 heterodimer from gastric mucosa.

Immunoprecipitation and Purification of the TFF1 Heterodimer. The TFF1 heterodimer was detected in TFF1 immunoprecipitates, and after reduction of the immunoprecipitates, the TFF1 immunoreactivity migrated in the same position as the 6.65 kDa recombinant TFF1. The TFF1 antibody does not react with recombinant TFF2 and TFF3

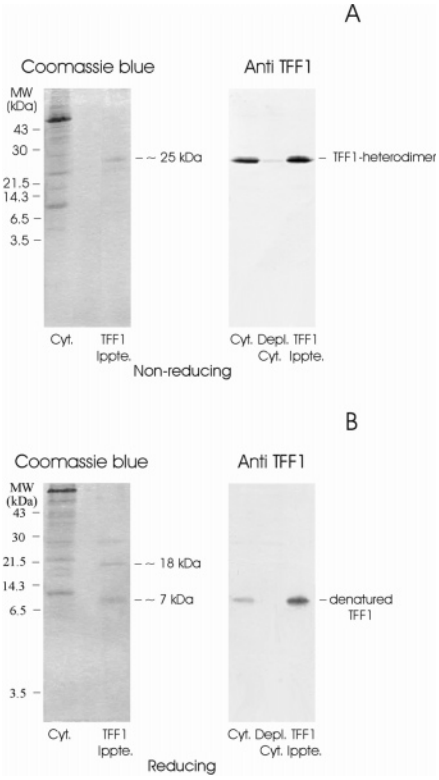


FIGURE 2: Purification of the TFF1 heterodimer by immunoprecipitation. Aliquots of gastric cytosol and of TFF1 immunoprecipitated from cytosol using purified antibody cross-linked to Sepharose beads were electrophoresed on polyacrylamide gels without (A) or after prior incubation with β-mercaptoethanol (B). The proteins were stained with Coomassie blue or transferred to PVDF membrane and incubated with anti-TFF1 antibodies. Aliquots of immunodepleted cytosol were included in the gels for immunodetection with anti-TFF1. The positions of the molecular mass markers are shown on the left and of the stained protein bands and different forms of TFF1 detected on the right of the appropriate panels.

that were included on the gels nor do TFF2 and TFF3 specific antibodies react with the TFF1 immunoprecipitate (data not shown). The absence of a reaction confirms the specificity of the TFF1 antibody and that TFF2 and TFF3 are not present in the TFF1 heterodimer.

Identification of the TFF1 protein partner in the heterodimer by amino acid sequencing and mass spectroscopy necessitated larger scale immunoprecipitation. TFF1 monoclonal antibody was purified and conjugated to immobilized protein G. The TFF1 heterodimer was purified from gastric cytosol, electrophoresed under nonreducing conditions, and stained with Coomassie blue (Figure 2A). The immunoprecipitate comprised a major protein band of ~25 kDa that coelectrophoresed with TFF1 heterodimer and which was not discernible among the proteins in complete cytosol (Figure 2A). Treatment with β-mercaptoethanol converted the majority of the ~25 kDa protein into two proteins of ~18 and ~7 kDa. Anti-TFF1 antiserum reacted with the ~7 kDa protein band only (Figure 2B). Analysis of the proteins that remained in gastric cytosol following TFF1 immunoprecipitation showed that essentially all the TFF1 heterodimer and TFF1 were removed by the immunoprecipitation (Depl. Cyt. in Figure 2A,B). The results show that the TFF1 complex is a heterodimer between an ~18 kDa protein and TFF1 and that the proteins are linked by a disulfide bond.

Identification of the TFF1 Protein Partner

Protein Sequencing. The TFF1 heterodimer was immunoprecipitated from gastric cytosol, treated with β-mercaptoethanol, electrophoresed on a polyacrylamide gel, and transferred to PVDF membrane. The proteins on the filter were stained with Coomassie blue, and the prominent band of ~18 kDa was excised and sequenced (Figure 3). An amino-terminal sequence of 13 amino acids, YEVFNIISPSNNG, was obtained. Searching the protein databases identified a 184 amino acid protein, XP114236, that had been predicted from the sequence of the human genome. The 13 amino acids correspond to residues 21−33, which shows that the first 20 residues comprise a signal sequence that is removed during synthesis of the protein partner.

Confirmation that XP114236 is the protein partner in the TFF1 heterodimer was sought by amino-terminal sequencing of the proteins present in the nonreduced TFF1 heterodimer. TFF1 was immunoprecipitated from a different gastric cytosol, electrophoresed on a nonreducing gel, and transferred to PVDF membrane. Amino-terminal sequencing of the proteins present in the nonreduced TFF1 heterodimer showed that XP114236 is the protein partner in the TFF1 heterodimer. Both XP114236 and TFF1 were present in the ~25 kDa protein band. Sequencing of the ~18 and ~7 kDa bands obtained after reduction of the TFF1 immunoprecipitate from the second gastric cytosol (Figure 3B) identified XP114236 as the larger and TFF1 as the smaller protein.

Mass Spectrometry. The protein sequencing results were corroborated by mass spectrometry of tryptic digests. Figure 4 shows the predicted trypsin cleavage sites in secreted XP114236 and TFF1 and the masses of the tryptic peptides that would be generated. Immunoprecipitated TFF1 heterodimer was excised from a nonreducing polyacrylamide gel and digested with trypsin. Four of the predicted trypsin cleavage products of secreted XP114236 were identified (Figure 4). This result was confirmed with a second gastric sample.

None of the tryptic peptides identified contain cysteine residues. Disulfide bonds could restrict trypsin access, or fragments could be joined. We therefore repeated the experiment with the TFF1 heterodimer that had been reduced and alkylated prior to electrophoresis. Five tryptic fragments of XP114236 were identified in the ~18 kDa protein band, one of which was identified also in the negative ion mode. One peptide resulted from incomplete trypsin digestion. The four amino-terminal tryptic peptides of TFF1 were identified in the ~7 kDa protein in the positive ion mode. Two peptides resulted from incomplete trypsin digestion. The carboxy-terminal tryptic peptide of TFF1 was identified in the negative ion mode. The mass spectrometry experiments identified ~50% (49%) of the XP114236 sequence and the entire TFF1 sequence in tryptic digests of TFF1 immunoprecipitates.

The amino-terminal sequencing and MALDI TOF mass spectrometry show that the TFF1 complex comprises a heterodimer between XP114236 and TFF1. We have called the TFF1 protein partner TFIZ1 for *t*refoil *f*actor *i*nteractions(*z*) 1.

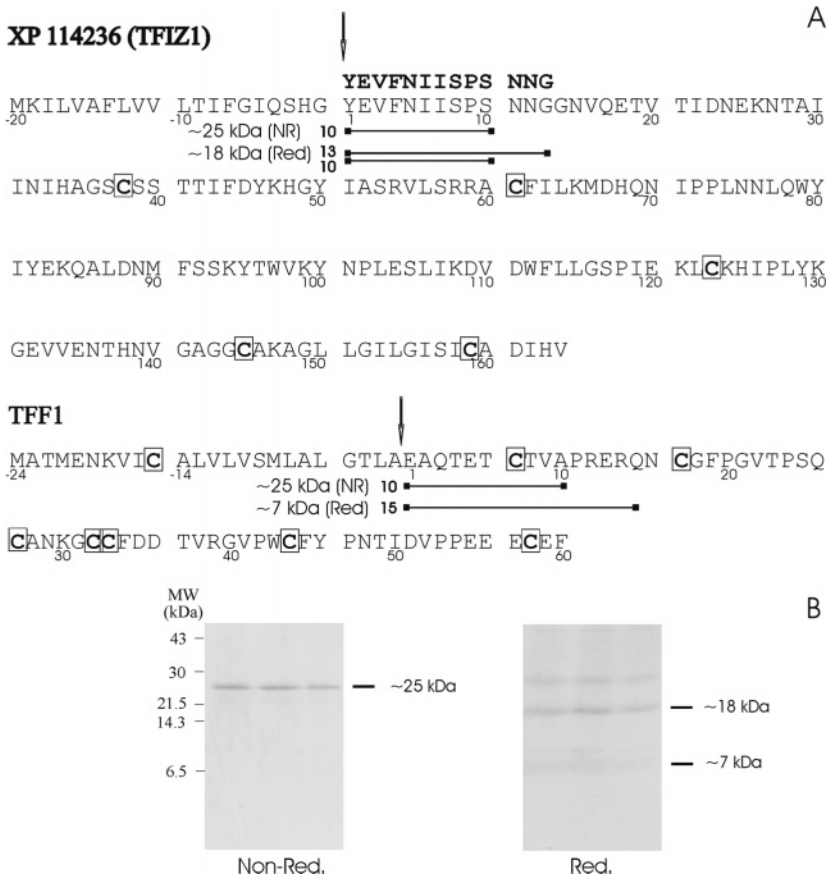


FIGURE 3: Identification of the protein partner in the TFF1 heterodimer by amino-terminal sequencing. (A) The amino acid sequences of XP114236 (TFIZ1) and of TFF1 are shown. The amino-terminal peptide sequence used to search the databases is shown in bold aligned above the corresponding XP114236 sequence. The sequences obtained from the ~25 kDa protein band are indicated below the corresponding XP114236 and TFF1 sequences (NR). The sequences derived from the ~18 kDa protein band are indicated below the XP114236 sequence (Red) and the sequence from the ~7 kDa protein is indicated below the TFF1 sequence (Red). The positions of the signal sequence splice sites are indicated by downward-pointing arrows, and the cysteine residues are shown in bold and are boxed. (B) Aliquots of TFF1 immunoprecipitated from cytosol using purified antibody cross-linked to Sepharose beads were electrophoresed on polyacrylamide gels without (Non-Red.) or after prior incubation with β-mercaptoethanol (Red.). The proteins were transferred to PVDF membrane and stained with Coomassie blue.

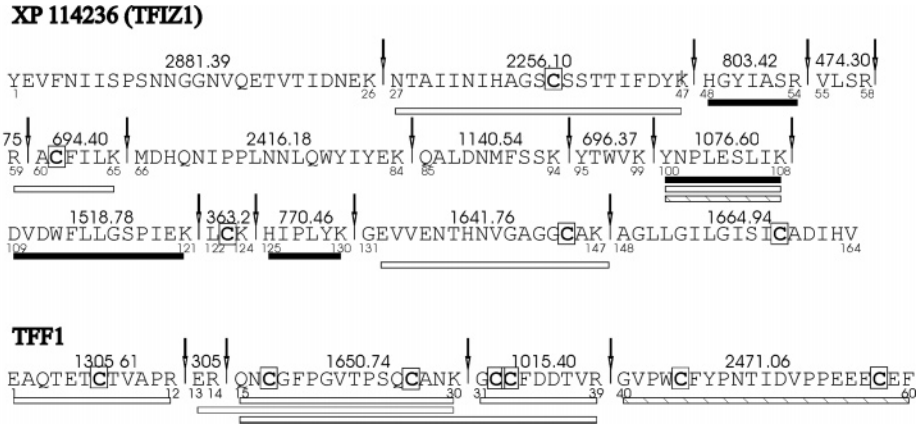


FIGURE 4: Identification of the protein partner in the TFF1 heterodimer by mass spectrometry of tryptic digests. The amino acid sequences of mature XP114236 (TFIZ1) (a) and TFF1 (b) are shown with the positions of predicted tryptic cleavage sites indicated by downward-pointing arrows. The masses of the cleavage fragments that would be generated are shown above the sequence. The fragments identified by MALDI TOF of tryptic digests of the ~25 kDa protein band are indicated by a solid line below the sequences. The fragments identified by MALDI TOF of tryptic digests of the ~18 or ~7 kDa bands are indicated below the sequences by an unfilled line for the positive ion mode and a hatch-filled line for the negative ion mode. The cysteine residues are shown in bold and are boxed.

Coexpression of TFF1 and TFIZ1 in Human Gastric Mucosa

Expression of TFIZ1. The *TFIZ1* gene is located on chromosome 2p13, 23.7 kb upstream from the *CA11/Amp-18* gene that encodes a gastric-specific protein (*29, 30*). Sequence comparison indicates that the two genes are related, and the possibility therefore exists that CA11/Amp-18 interacts with trefoil proteins and should be designated

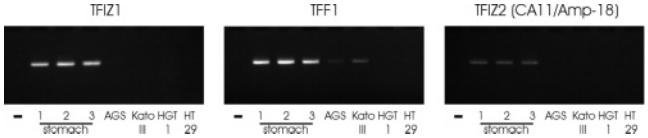


FIGURE 5: Expression of gastric TFIZ1 and TFIZ2. RNA was prepared from samples of normal stomach (1−3) and from AGS, Kato-III, HGT-1, and HT29 cells. The RNA was reverse transcribed and amplified by PCR with primer pairs for TFIZ1, TFF1, and TFIZ2 (CA11/AMP-18). The products were separated by electrophoresis on a 3% agarose gel and stained with ethidium bromide.

TFIZ2. The predicted TFIZ1 sequence, TFF1, and TFIZ2 (CA11/AMP-18) mRNAs were amplified by RT PCR from human gastric RNA and from three gastric type and HT29 intestinal cell line RNAs (Figure 5). PCR products were detected from normal stomach for TFIZ1, TFF1, and TFIZ2 (CA11/AMP-18) mRNAs. PCR products were detected, in smaller amounts, for TFF1 but not TFIZ1 or TFIZ2 (CA11/AMP-18) mRNAs after amplification of AGS, Kato III, and HT29 cell RNA. This suggested that TFF1 may be expressed in the absence of its normal protein partner TFIZ1 in gastric tumor cells.

TFIZ1 Sequence. Human gastric TFIZ1 mRNA was sequenced and compared to the sequence predicted from the human genome and the sequence of human TFIZ2. Full-length TFIZ1 and TFIZ2 (CA11/AMP-18) mRNAs were amplified from gastric mucosal RNA and inserted into the vector pCR2.1 by TOPO cloning and sequenced (Figure 6). The TFIZ1 sequence obtained from several individuals is identical to the XP114236 mRNA sequence deduced from the human genome. There are two differences between the TFIZ2 sequence that we obtained and the published sequence (*29*). In exon 2, the codon for leucine that is located two amino acids terminal to the signal peptide cleavage site is CTA, and there is an A in place of a G at position 695 in exon 6. These two differences were identified in TFIZ2 isolated from several individuals. Both genes contain six exons. Exon 1 of TFIZ2 is 75 nucleotides longer than that of TFIZ1, and exons 2, 3, and 4 are the same lengths in both genes whereas the lengths of exons 5 and 6 differ slightly. Overall nucleotide similarity between the genes is 42% but varies throughout the sequences and is most similar between the third exons (57%).

Characteristics of the TFIZ1 Protein. The TFIZ1 mRNA sequence contains an open reading frame that encodes a protein of 184 amino acids. We have sequenced the amino terminus of TFIZ1, and this demonstrates that mature TFIZ1 is a 164 amino acid secreted protein with a theoretical molecular mass of 18.31 kDa and an isoelectric point of 6.59. The TFF1 heterodimer has a predicted molecular mass of 24.98 kDa, which is remarkably close to the apparent molecular mass of ~25 kDa. The theoretical isoelectric point of the heterodimer is 5.16. TFIZ1 has five cysteine residues and will have therefore a free cysteine available to interact with TFF1 Cys58 (Figure 6).

The protein database was searched with TFIZ1, which showed that TFIZ1 contains two recognized conserved homology domains (Figure 7A). The first is smart00019, SF_P, which is found in the pulmonary surfactant-associated protein C precursor. TFIZ1 shares homology with the carboxy-terminal end of smart00019, SF_P, with a bit score of 41.2. The second homology is with the brichos domain.

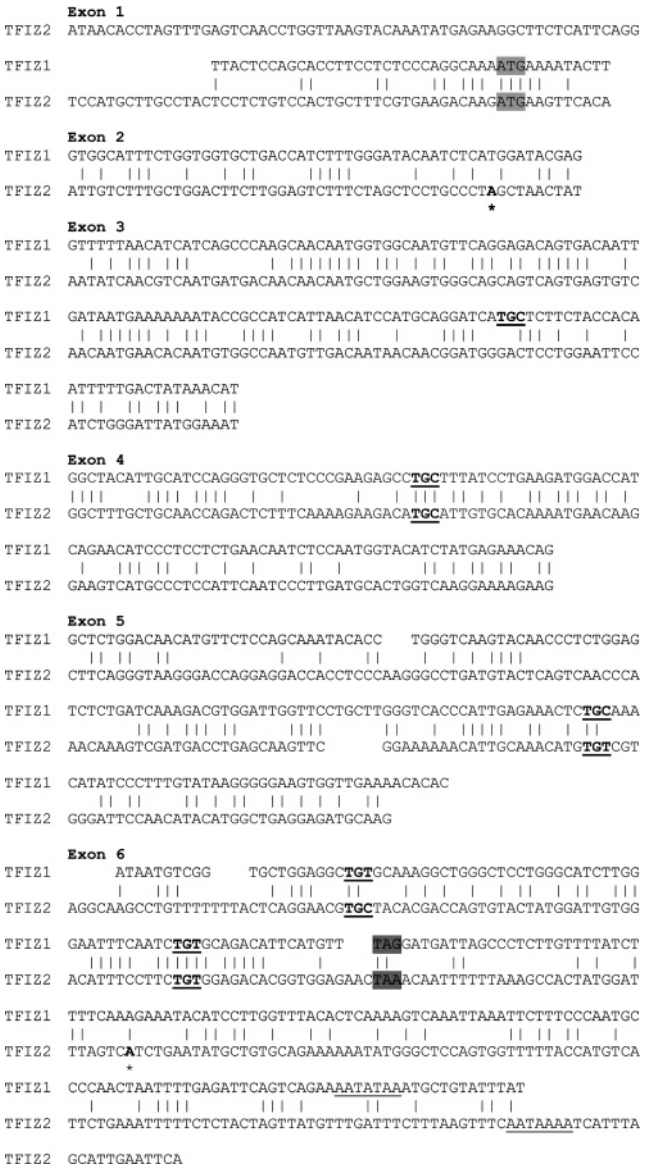
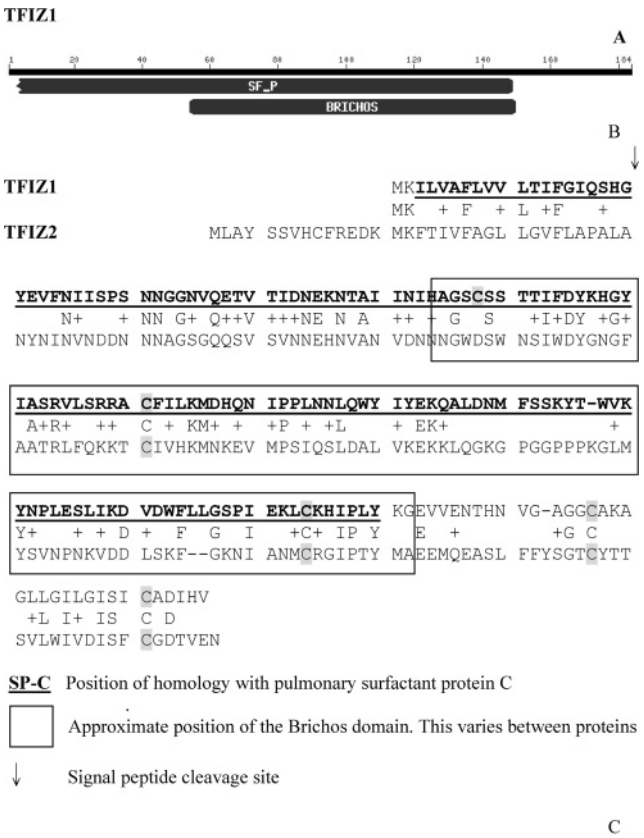


FIGURE 6: Comparison of the TFIZ1 and TFIZ2 cDNA sequences. The sequences of TFIZ1 and TFIZ2 cDNAs are shown aligned. The nucleotides in TFIZ2 that differ from the published sequences of CA11 are indicated by an asterisk and shown in bold. The nucleotides of the initiation methionine (pale gray) and stop (gray) codons are highlighted. The nucleotides that encode the cysteine residues are underlined and shown in bold, and the transcription termination signals are underlined.

The brichos domain is about 100 amino acids long and is of unknown function. TFIZ1 shares homology with the entire brichos domain with a bit score of 87.3. Three of the five TFIZ1 cysteine residues occur within the brichos domain, and of these Cys38 is found only in TFIZ1 whereas Cys61 and Cys123 are conserved between all brichos domain-containing proteins. This suggests that the intermolecular disulfide bond in the TFF1 heterodimer may be between Cys38 of TFIZ1 and Cys58 of TFF1 (Figure 8).

The sequences of the TFIZ1 and TFIZ2 proteins are shown aligned in Figure 7B. TFIZ1 lacks amino acids Lys97 and Tyr141 of TFIZ2 (CA11/AMP-18), and the protein is terminated one residue earlier. TFIZ2 (CA11/AMP-18) does not contain Leu114 and Leu115 of TFIZ1. Four of the five TFIZ1 cysteine residues are conserved between the two proteins, but Cys38 is unique to TFIZ1. This reinforces the



C

Distribution of conservation between TFIZ1 and TFIZ2 (CA11/AMP-18)

	Identical	Semi-conserved
Preprotein	26 % (48/184)	49 % (90/184)
Mature protein	26 % (43/164)	49 % (81/164)
N-terminal domain	30 % (10/33)	61 % (20/33)
Brichos domain	24 % (24/99)	48 % (48/99)
C-terminal domain	26 % (9/34)	38 % (13/34)

FIGURE 7: The TFIZ family of proteins. (A) Positions of homology with smart00019, SF_P (SF_P), and the brichos (BRICHOS) domain in the TFIZ1 preprotein. (B) Amino acid sequences for TFIZ1 and TFIZ2 are shown aligned. Conserved residues are listed, and semiconserved residues are indicated by a (+) sign between the two sequences. The region of homology between TFIZ1 and smart00019, SF_P, is shown in bold and underlined. The brichos domain is boxed. A downward-pointing arrow indicates the positions of the signal sequence cleavage sites. The cysteine residues are highlighted (gray). (C) The conservation between TFIZ1 and TFIZ2 is listed.

notion that the intermolecular disulfide bond in the TFF1:TFIZ1 heterodimer involves Cys38 of TFIZ1.

Amino acid residues that are identical or conserved between TFIZ1 and TFIZ2 are distributed relatively evenly throughout the sequences (Figure 7C). Conservation is higher at the amino termini encoded by the third exons. These regions of the proteins may have important functional roles. There is a compositionally biased, low complexity sequence surrounding Cys145 in TFIZ1 which comprises 6 glycine, 3 alanine, 3 leucine, and 3 isoleucine residues within a stretch of 18 amino acid residues (Figure 7B). None of the alanine residues and only one each of the glycine, leucine, and isoleucine residues are conserved in TFIZ2, which indicates that the function of this region of the protein is unique to TFIZ1.

Interaction between TFF1 and TFIZ1. Recombinant TFF1 forms homodimers, and TFF1 dimers have been detected in gastric mucosa and shown to interact with MUC5AC (*8*). The TFF1 heterodimer could result from an interaction between a TFF1 homodimer and another protein. The demonstration that the TFF1:TFIZ1 heterodimer is resistant to treatment with ionic detergents and boiling but is destroyed by incubation in the presence of a thiol agent argues against this possibility (Figure 2). We have shown previously that the heterodimer is resistant to treatment with guanidine hydrochloride and have demonstrated that it is resistant to treatment with urea and nonionic detergents and to proteolysis (unpublished data). This indicates that the most important bond stabilizing the interaction between TFF1 and its protein partner is the covalent disulfide bond between the two molecules.

The apparent molecular masses of TFF1 and TFIZ1 and of the TFF1:TFIZ1 heterodimer (Figure 2) are also consistent with the proposition that one molecule of TFF1 interacts directly via a disulfide bond with one molecule of TFIZ1 to form the heterodimer. TFIZ1 Cys38, which is the cysteine residue most likely to form the disulfide bond with Cys58 of TFF1, is located within the brichos domain (Figure 8). This indicates that the brichos domain is involved in the interaction between TFF1 and TFIZ1. In contrast, Cys58 lies outside the trefoil domain of TFF1. The concentration of acidic residues that surround Cys58 at the carboxy terminus of TFF1 has been noted previously (Figure 1A; *6*, *7*, *22*). Interestingly, TFIZ1 Cys38 is embedded within a stretch of polar amino acid residues: three serines and two threonines.

DISCUSSION

Identification of the TFF1:TFIZ1 Heterodimer. We have purified the most abundant molecular form of TFF1 in human gastric mucosa and shown that it is a heterodimer between TFF1 and a previously unknown protein that we have called TFIZ1. We have sequenced TFIZ1 mRNA from human gastric mucosa and have shown that TFIZ1 is a secreted protein and that the TFIZ1 monomer has a theoretical molecular mass of 18.3 kDA. TFIZ1 and TFF1 are coexpressed in the normal human stomach. TFIZ1 expression was not detected in gastric cancer cells, but TFF1 expression was detected in some gastric cancer cells.

Role of the Heterodimer in TFF1 Biology. The TFF1:TFIZ1 heterodimer is the major form of TFF1 in normal human gastric mucosa, which indicates that its formation is critical to TFF1 function. The heterodimer could have enhanced activity relative to TFF1 or have biologically distinct functions from TFF1. Alternatively, formation of the heterodimer could prolong the half-life of TFF1, which might be important in the harsh environment of the stomach. Another possibility is that TFIZ1 sequesters free TFF1 and thereby regulates or reduces its action.

Dissociation of TFF1 and TFIZ1 expression in gastric carcinoma cells could provide an explanation for the apparent contradiction that TFF1 expression is beneficial in normal tissue but detrimental in cancerous tissue (*6*). It may also reconcile the observation that a gastric tumor suppressor is expressed ectopically by other tumors and at high levels by a large proportion of breast tumors including interval cancers (*31*−*34*). TFIZ1 may itself be a tumor suppressor, and the reduced expression of TFIZ1 in human gastric tumor cells is supportive of this possibility.

Homology between TFIZ1 and Other Proteins. The *TFIZ1* gene is located on chromosome 2p13.3 and is transcribed

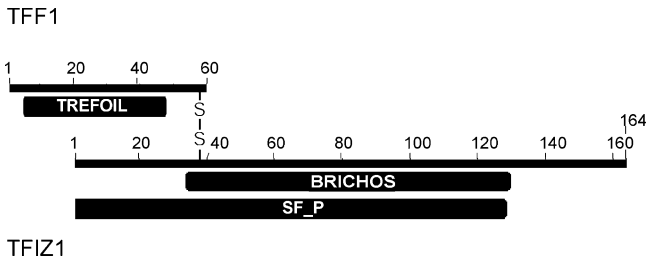


FIGURE 8: Schematic representation of the interaction between TFF1 and TFIZ1 in the heterodimer. The positions of the conserved domains present in the proteins are indicated as gray bars. These are the trefoil domain (TREFOIL) in TFF1 and the brichos domain (BRICHOS) and smart00019, SF_P domain (SF_P) of pulmonary surfactant-associated protein C precursor in TFIZ1. The disulfide bond between TFF1 Cys58 and TFIZ1 Cys38 is shown.

from the centromere to the telomere. The *TFIZ2* (*CA11/AMP-18*) gene (*29*, *30*) is 23.7 kb upstream and is transcribed in the opposite direction to the *TFIZ1* gene.

TFIZ1 shares homology with the pulmonary surfactant-associated protein C precursor domain, smart00019, SF_P (Figure 7A). This region of homology is not identified in TFIZ2. Pulmonary surfactant protein C helps to maintain cellular function at the interface between the alveoli surface and the airways (*35*). The expression of TFIZ1 in the gastric mucosa indicates that it functions at the interface between the gastrointestinal mucosa and lumen. It seems probable that the shared homology between pulmonary surfactant-associated protein C precursor and TFIZ1 is indicative of some shared functional attributes. The region of homology is within the proprotein of surfactant protein C which is thought to be involved in vesicular trafficking.

TFIZ1 contains one copy of the conserved brichos domain (Figure 7) (*36*). The brichos domain is about 100 amino acids long and therefore comprises over 50% of the TFIZ1 protein sequence. The brichos domain is present in proteins associated with dementia, respiratory distress, and cancer. Our demonstration that the mucosal protective trefoil protein TFF1 is covalently bound to TFIZ1 links these two protein families. This is the first demonstration that a member of the trefoil factor family of proteins is bound to a brichos domain-containing protein. It will be interesting to determine if TFIZ2 interacts with TFF1 and if either of the two proteins interacts with TFF2 or TFF3.

Interaction between TFF1 and TFIZ1. TFF1 is a very acidic protein with a p*I* of 3.94 (*22*), whereas TFIZ1 is neutral with a p*I* of 6.59. The p*I* of the TFF1:TFIZ1 heterodimer is still relatively acidic at 5.16. The theoretical molecular mass of the TFF1:TFIZ1 heterodimer is 24.98 kDa, which agrees very well with the molecular mass determined empirically. Our data show therefore that a single monomer unit of TFF1 binds directly to TFIZ1 most probably via a disulfide bond between Cys58 of TFF1 and Cys38 of TFIZ1 to form a stable heterodimer (Figure 8). Cys38 is located within the brichos domain, and this shows that the region of direct contact between the surfaces of TFIZ1 and TFF1 includes a portion of the brichos domain.

ACKNOWLEDGMENT

We thank Dr. Joe Gray and Mr. Bob Liddell for protein and DNA sequencing and for performing MALDI TOF mass spectrometry.

REFERENCES

1. Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A., and Chambon, P. (1982) Cloning of cDNA sequences of hormone-regulated genes from the MCF-7 human breast cancer cell line, *Nucleic Acids Res. 10*, 7895−7905.
2. May, F. E. B., and Westley, B. R. (1986) Cloning of estrogen-regulated messenger RNA sequences from human breast cancer cells, *Cancer Res. 46*, 6034−6040.
3. Rio, M. C., Bellocq, J. P., Daniel, J. Y., Tomasetto, C., Lathe, R., Chenard, M. P., Batzenschlager, A., and Chambon, P. (1988) Breast cancer-associated pS2 protein: synthesis and secretion by normal stomach mucosa, *Science 241*, 705−708.
4. Piggott, N. H., Henry, J. A., May, F. E. B., and Westley, B. R. (1991) Antipeptide antibodies against the pNR-2 oestrogen-regulated protein of human breast cancer cells and detection of pNR-2 expression in normal tissues by immunohistochemistry, *J. Pathol. 163*, 95−104.
5. Poulsom, R., and Wright, N. A. (1993) Trefoil peptides: a newly recognized family of epithelial mucin-associated molecules, *Am. J. Physiol. 265*, G205−G213.
6. May, F. E. B., and Westley, B. R. (1997) Trefoil proteins: their role in normal and malignant cells, *J. Pathol. 183*, 4−7.
7. Ribieras, S., Tomasetto, C., and Rio, M. C. (1998) The pS2/TFF1 trefoil factor, from basic research to clinical applications, *Biochim. Biophys. Acta 1378*, F61−F77.
8. Ruchaud-Sparagano, M.-H., Westley B. R., and May, F. E. B. (2004) The trefoil protein TFF1 is bound to MUC5AC in human gastric mucosa, *Cell. Mol. Life Sci. 61*, 1946−1954.
9. Clyne, M., Dillon, P., Daly, S., O'Kennedy, R., May, F. E. B., Westley, B. R., and Drumm, B. (2004) *Helicobacter pylori* interacts with the human single-domain trefoil protein TFF1, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7409−7414.
10. Playford, R. J., Marchbank, T., Goodlad, R. A., Chinery, R. A., Poulsom, R., and Hanby, A. M. (1996) Transgenic mice that overexpress the human trefoil peptide pS2 have an increased resistance to intestinal damage, *Proc. Natl. Acad. Sci. U.S.A. 93*, 2137−2142.
11. Marchbank, T., Westley, B. R., May, F. E. B., Calnan, D. P., and Playford R. J. (1998) Dimerization of human pS2 (TFF1) plays a key role in its protective/healing effects, *J. Pathol. 185*, 153−158.
12. Lefebvre, O., Chenard, M.-P., Masson, R., Linares, J., Dierich, A., LeMeur, M., Wendling, C., Tomasetto, C., Chambon P., and Rio, M.-C. (1996) Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein, *Science 274*, 259−262.
13. Henry, J. A., Bennett, M. K., Piggott, N. H., May, F. E. B., and Westley, B. R. (1991) Expression of pNR-2/pS2 protein in diverse human epithelial tumours, *Br. J. Cancer 64*, 677−682.
14. Muller, W., and Borchard, F. (1993) pS2 protein in gastric carcinoma and normal gastric mucosa: association with clinico-pathological parameters and patient survival, *J. Pathol. 171*, 263−269.
15. Park, W. S., Oh, R. R., Park, J. Y., Lee, J. H., Shin, M. S., Kim, H. S., Lee, H. K., Kim, Y. S., Kim, S. Y., Lee, S. H., Yoo, N. J., and Lee, J. Y. (2000) Somatic mutations of the trefoil factor family 1 gene in gastric cancer, *Gastroenterology 119*, 691−698.
16. Park, W. S., Oh, R. R., Park, J. Y., Yoo, N. J., Lee, S. H., Shin, M. S., Kim, S. Y., Kim, Y. S., Lee, J. H., Kim, H. S., An, W. G., and Lee, J. Y. (2000) Mapping of a new target region of allelic loss at 21q22 in primary gastric cancers, *Cancer Lett. 159*, 15−21.
17. Fujimoto, J., Yasui, W., Tahara, H., Tahara, E., Kudo, Y., Yokozaki, H., and Tahara, E. (2000) DNA hypermethylation at the pS2 promoter region is associated with early stage of stomach carcinogenesis, *Cancer Lett. 149*, 125−134.
18. Prest, S. J., May, F. E. B., and Westley, B. R. (2002) The estrogen-regulated protein, TFF1, stimulates migration of human breast cancer cells, *FASEB J. 16*, 592−594.
19. Emami, S., Le Floch, N., Bruyneel, E., Thim, L., May, F., Westley, B., Rio, M., Mareel, M., and Gespach, C. (2001) Induction of scattering and cellular invasion by trefoil peptides in src- and RhoA-transformed kidney and colonic epithelial cells, *FASEB J. 15*, 351−361.
20. Bossenmeyer-Pourie, C., Kannan, R., Ribieras, S., Wendling, C., Stoll, I., Thim, L., Tomasetto, C., and Rio, M. C. (2002) The trefoil factor 1 participates in gastrointestinal cell differentiation by delaying G1-S phase transition and reducing apoptosis, *J. Cell Biol. 157*, 761−770.

21. Rodrigues, S., Van Aken, E., Van Bocxlaer, S., Attoub, S., Nguyen, Q. D., Bruyneel, E., Westley, B. R., May, F. E. B., Thim, L., Mareel, M., Gespach, C., and Emami, S. (2003) Trefoil peptides as proangiogenic factors in vivo and in vitro: implication of cyclooxygenase-2 and EGF receptor signalling, *FASEB J. 17*, 7−16.
22. May, F. E. B., Church, S. T., Major, S., and Westley, B. R. (2003) The closely related estrogen-regulated trefoil proteins TFF1 and TFF3 have markedly different hydrodynamic properties, overall charge, and distribution of surface charge, *Biochemistry 42*, 8250−8259.
23. Chadwick, M. P., Westley, B. R., and May, F. E. B. (1997) Homodimerization and hetero-oligomerization of the single-domain trefoil protein pNR-2/pS2 through cysteine, *Biochem. J. 327*, 117−123.
24. Polshakov, V. I., Williams, M. A., Gargaro, A. R., Frenkiel, T. A., Westley, B. R., Chadwick, M. P., May, F. E. B., and Feeney, J. (1997) High-resolution structure of the human pNR-2/pS2 single domain trefoil peptide, *J. Mol. Biol. 267*, 418−432.
25. Williams, M. A., Westley, B. R., May, F. E. B., and Feeney, J. (2001) The solution structure of the disulphide-linked homodimer of the human trefoil protein TFF1, *FEBS Lett. 493*, 70−74.
26. Newton, J. L., Allen, A., Westley, B. R., and May, F. E. B. (2000) The human trefoil peptide, TFF1, is present in different molecular forms that are intimately associated with mucus in normal stomach, *Gut 46*, 312−320.
27. Chadwick, M. P., May, F. E. B., and Westley, B. R. (1995) Production and comparison of mature single-domain "trefoil" peptides pNR-2/pS2 Cys^{58} and pNR-2/pS2 Ser^{58}, *Biochem. J. 308*, 1001−1007.
28. Carr, M., May, F. E. B., Lennard, T. W. J., and Westley, B. R. (1995) Determination of oestrogen responsiveness of breast cancer by competitive reverse transcription-polymerase chain reaction, *Br. J. Cancer 72*, 1427−1434.
29. Yoshikawa, Y., Mukai, H., Hino, F., Asada, K., and Kato, I. (2000) Isolation of two novel genes, down-regulated in gastric cancer, *Jpn. J. Cancer Res. 91*, 459−463.
30. Martin, T. E., Powell, C. T., Wang, Z., Bhattacharyya, S., Walsh-Reitz, M. M., Agarwal, K., and Toback, F. G. (2003) A novel mitogenic protein that is highly expressed in cells of the gastric antrum mucosa, *Am. J. Physiol.: Gastrointest. Liver Physiol. 285*, G332−G343.
31. Rio, M. C., Bellocq, J. P., Gairard, B., Rasmussen, U. B., Krust, A., Koehl, C., Calderoli, H., Schiff, V., Renaud, R., and Chambon, P. (1987) Specific expression of the pS2 gene in subclasses of breast cancers in comparison with expression of the estrogen and progesterone receptors and the oncogene, *Proc. Natl. Acad. Sci. U.S.A. 84*, 9243−9247.
32. Henry, J. A., Piggott, N. H., Mallick, U.K., Nicholson, S., Farndon, J. P., Westley, B. R., and May, F. E. B. (1991) pNR-2/pS2 immunohistochemical staining in breast cancer: correlation with prognostic factors and endocrine response, *Br. J. Cancer 63*, 615−622.
33. Foekens, J. A., Portengen, H., Look, M. P., van Putten, W. L., Thirion, B., Bontenbal, M., and Klijn, J. G. (1994). Relationship of PS2 with response to tamoxifen therapy in patients with recurrent breast cancer, *Br. J. Cancer 70*, 1217−1223.
34. Crosier, M., Scott, D., Wilson, R. G., Griffiths, C. D., May, F. E. B., and Westley, B. R. (2001) High expression of the trefoil protein TFF1 in interval breast cancers, *Am. J. Pathol. 159*, 215−221.
35. Weaver, T. E., and Conkright, J. J. (2001) Functions of surfactant proteins B and C, *Annu. Rev. Physiol. 63*, 555−578.
36. Sanchez-Pulido, L., Devos, D., and Valencia, A. (2002) BRICHOS: a conserved domain in proteins associated with dementia, respiratory distress and cancer, *Trends Biochem. Sci. 27*, 329−332.

BI047287N