The Closely Related Estrogen-Regulated Trefoil Proteins TFF1 and TFF3 Have Markedly Different Hydrodynamic Properties, Overall Charge, and Distribution of Surface Charge†

Felicity E. B. May, Stephen T. Church, Sarah Major, and Bruce R. Westley*

Department of Pathology, School of Clinical and Laboratory Sciences, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP, U.K.

Received January 28, 2003

ABSTRACT: The human trefoil proteins TFF1 and TFF3 are expressed predominantly in the gastrointestinal tract. They are also expressed and regulated by estrogens in malignant breast epithelial cells. TFF1 and TFF3 are small cysteine-rich acidic secreted proteins of 60 and 59 amino acids with similar isoelectric points of 4.75 and 3.94, respectively. Each contains one trefoil domain that is characterized by several conserved features including six cysteine residues with conserved spacing. TFF1 and TFF3 form intermolecular disulfide bonds via an extra-trefoil domain cysteine residue and are present in vivo as monomers and homodimers and as complexes with other proteins. The TFF1 dimer is more active than the TFF1 monomer. In the present study the hydrodynamic and charge properties of TFF1 and TFF3 monomers and homodimers have been compared and shown to differ markedly. Notably, TFF1 is significantly more asymmetric than TFF3 (frictional coefficients 1.25 and 1.12, respectively, $p < 0.001$), and homodimerization of TFF1 results in a greater increase in asymmetry than for TFF3. The overall charges of TFF1 and TFF3 are very different at neutral pH. Titration curves predicted significant differences in charge across a wide pH range that agreed well with experimental data. The locations of charged amino acids in the primary sequences and in the tertiary structures of TFF1 and TFF3 were examined. This revealed interesting divergence in both the distribution and local topology of charged amino acid side chains. The significant differences between the shape, size, and surface charge of these two closely related molecules may account for their divergent biological activities.

It has long been recognized that breast cancer can be estrogen-responsive, and hormonal therapy is used successfully in its treatment. The identification of genes whose expression is regulated by estrogens in breast tumors should suggest mechanisms by which estrogen exerts its effects and may provide clinically useful markers of estrogen-responsive breast cancer. The most frequently identified estrogen-responsive gene encodes TFF1[1] (*1*, *2*). The expression of the closely related protein, TFF3, is also regulated by estrogen in breast cancer cells (*3*). The close association between estrogen receptor expression and TFF1 and TFF3 expression in breast cancer has been confirmed in two recent microarray-based studies of primary breast tumors (*4*, *5*). The value of TFF1 for prediction of response to hormonal therapy has been demonstrated (*6*).

The predominant site of expression of all three trefoil peptides in normal tissue is the gastrointestinal tract. Both TFF1 and TFF2 are expressed mainly in the stomach whereas TFF3 is expressed mainly in the small and large intestines (*7*, *8*). It is thought that the normal function of trefoil proteins is to protect and maintain the integrity of the gastrointestinal mucosa both by stabilizing the mucus barrier and by stimulating restitution. The role of TFF1 and TFF3 in breast cancer is uncertain, but the demonstration that TFF1 stimulates breast cancer cell migration suggests that it may promote tumor invasiveness (*9*).

TFF1 and TFF3 are small cysteine-rich secreted proteins of 60 and 59 amino acids, respectively. Both are acidic proteins with isoelectric points of 4.75 and 3.94, respectively, and are members of the trefoil factor family (TFF) of proteins that share considerable homology within the trefoil domain of 42 to 43 amino acids. The trefoil domain is characterized by several well-conserved features including six half-cystines with conserved spacing. There are three human trefoil peptides: TFF1 and TFF3 contain one trefoil domain while TFF2 contains two trefoil domains in a single chain of 106 amino acids. TFF2 is not expressed in breast cancer cells.

TFF1 and TFF3 both contain an extra-trefoil domain cysteine located three residues from the carboxy terminus (see Figure 6A). This cysteine forms intermolecular disulfide bonds to produce homo- and heterodimers (*10*, *11*), and the homodimer has greater biological activity than the monomer (*9*). The three-dimensional structures of the human TFF1 monomer and dimer (*12*, *13*) and the human TFF3 monomer (*14*) have been determined. Within the trefoil domain, the half-cystines are paired 1−5, 2−4, and 3−6 (Figure 1A) to form the compact trefoil motif which comprises three closely

† This work was supported by Cancer Research U.K., The Wellcome Trust, and the Breast Cancer Campaign.

* To whom correspondence should be addressed.

[1] Abbreviations: DTT, dithiothreitol; SDS, sodium dodecyl sulfate; TFF, trefoil factor family; BSA, bovine serum albumin.

10.1021/bi030025l CCC: $25.00
Published on Web 06/17/2003

packed loops. The third loop is positioned between the first and second. There is a short α-helix in loop 2 that is packed against two antiparallel β-strands from loop 3. Several conserved hydrophobic residues comprising adjacent parts of loops 2 and 3 are juxtaposed to form a patch on the surface of the molecules. It has been proposed that this area of the trefoil domain could bind to a receptor (*12*). In TFF1 and TFF3, the amino and carboxy termini are closely associated and form extended structures. Apart from some β-type interactions between residues 3−7 and residues 47−51 in TFF1 (*12*) the N- and C-termini have a poorly defined conformation in NMR structures, indicating that they have significant mobility in solution.

TFF1 and TFF3 both protect the gastrointestinal tract against damage and stimulate cell migration; however, the precise physiological functions of TFF1 and TFF3 are not understood (*7*, *8*). Data from TFF1 and TFF3 knockout mice indicate that they may be divergent, implying that they may elicit their actions through distinct protein−protein interactions (*15*, *16*). Although the clinical importance of TFF1 and TFF3 expression in breast tumors is recognized, their biological roles remain to be determined. In the present study we have compared the hydrodynamic and charge characteristics of TFF1 and TFF3 monomers and dimers to investigate whether there are differences in conformational and charge properties of the molecules that may determine their respective biological activities.

EXPERIMENTAL PROCEDURES

Production of Recombinant TFF1 and TFF3. The expression plasmids for TFF1 Cys^{58} and TFF1 Ser^{58} have been described previously (*10*, *17*). Human TFF3 cDNA was amplified by PCR from a KSM13 plasmid that contained TFF3 cDNA (*3*). The 5′ PCR primer included the recognition site for the *Eco*RI restriction endonuclease, adjacent to 12 bases that encode an in-frame consensus four amino acid recognition sequence for factor Xa, followed by the first 16 bases of TFF3 cDNA that encode the first five amino acids of the mature protein: 5′ GCGAATTCGATCGAAGGGCGTGAGGAGTACGTGGGCC 3′. Factor Xa is a specific proteinase that cleaves immediately 3′ of its four amino acid recognition site. Two 3′ PCR primers were used. The first contains the antisense strand for the last five amino acids and the stop codon of TFF3, followed by the recognition site for the *Xba*I restriction endonuclease: 5′ GCTCTAGATCAGAAGGTGCATTCTGC 3′. In combination with the 5′ PCR primer, this PCR primer amplifies wild-type TFF3 Cys^{57}. The second 3′ PCR primer contained the antisense strand to code for the last seven amino acids of TFF3 plus one base, except that the codon for Cys^{57}, TGC, was replaced with TCC, which encodes a serine residue (Figure 1A). The recognition site for the *Xba*I restriction endonuclease follows the stop codon of TFF3: 5′ GCTCTAGATCAGAAGGTGGATTCTGCTTCCTGC 3′. In combination with the 5′ PCR primer, this PCR primer amplifies TFF3 Ser^{57}. The PCR products were ligated into the multiple cloning site of the pEZZ18 vector that had been digested previously with *Eco*RI and *Xba*I. The cloned DNA was sequenced in both directions to confirm that it contained the correct DNA sequence.

BL21 *Escherichia coli* cells that contained the recombinant plasmids were grown overnight to stationary phase at 37 °C, retrieved by centrifugation, and washed in 20% (w/v) sucrose, 1 mM EDTA, and 0.3 M Tris-HCl, pH 8.0 (*17*). The fusion protein was released from the periplasm by subjecting the cells to an osmotic shock in 0.5 mM $MgCl_2$ for 20 min at 4 °C and purified by affinity chromatography on a 40 mL IgG−Sepharose (Pharmacia) column equilibrated in 150 mM NaCl and 50 mM Tris-HCl, pH 7.6. Nonspecifically bound material was removed in 5 mM ammonium acetate, pH 5.0, and the bound fusion protein eluted with 0.5 M acetic acid, pH 3.4. The protein was transferred to 100 mM NaCl, 1 mM $CaCl_2$, and 50 mM Tris-HCl, pH 8.0, and digested for 48−72 h at 37 °C with factor Xa. The ZZ fusion partner was removed from the mature TFF3 protein by passage over IgG−Sepharose.

Purification of TFF1 and TFF3. Following the second passage over IgG−Sepharose, 50 mM cysteine was added to TFF1 Cys^{58} and 1 mM DTT to TFF3 Cys^{57}, and the proteins were incubated at 4 °C for 48 and 20 h, respectively. The recombinant proteins were bound to a 1 mL mono Q anion-exchange column in piperazine hydrochloride, pH 5, for TFF1 and Tris-HCl, pH 8.5, for TFF3 and eluted with NaCl. The proteins were concentrated using a Vivaspin concentrator (Vivascience) with a 5000 molecular weight cutoff filter. They were purified further by gel filtration on Superdex 75 (Pharmacia) in 50 mM sodium phosphate buffer, pH 6.5, and concentrated to 2 mg/mL with a Vivaspin concentrator. Protein concentrations were measured using the bicinchoninic acid assay (Pierce) with BSA as the standard. The proteins were stored at −20 °C.

Polyacrylamide Gel Electrophoresis. The separating gels shown in Figure 1B,C contained 20% acrylamide (w/v) and the stacking gel 10% (w/v) acrylamide; both contained 10% (w/v) glycerol. The separating gels shown in Figure 4 contained a gradient of 12−32% (w/v) acrylamide and the stacking gels 5% (w/v) acrylamide (*10*). There was no SDS or reducing agent in any of the buffers used in the gels shown in Figures 1B,C and 4B, but for the gels shown in Figure 4A, 0.1% SDS was included in all of the buffers. The proteins were brought to 62.5 mM Tris-HCl, pH 6.8, 12.5 mM EDTA, 10% glycerol, and 0.005% bromophenol blue and were not heated before being loaded for the gels shown in Figures 1B and 4B. For the gels shown in Figure 4A, 2% SDS and 100 mM β-mercaptoethanol were included in the loading buffer, and the samples were boiled for 5 min before being loaded. The gels were stained as described previously (*17*).

Characterization of TFF1 and TFF3. Molecular masses under native conditions were measured by equilibrium ultracentrifugation on a Beckman XLA instrument at 20 °C at three protein concentrations (1.4, 1.1, and 0.7 mg/mL) and three different rotor speeds. The partial specific volume was calculated from the amino acid composition at 0.706 mL/g and assumed to be the same in both monomeric and dimeric states. The ideal noninteracting model was adequate to describe all of the data.

Stokes radii were measured by gel filtration on a Superdex 75 column in 50 mM sodium phosphate buffer, pH 7.0, and 150 mM NaCl. The column was calibrated with aprotinin, RNase A, ovalbumin, BSA, and Blue Dextran.

Titration curves were computed for TFF1 and TFF3 on the basis of the amino acid compositions shown in Table 1.

For ion-exchange chromatography at different pH values, the proteins were diluted 10-fold into the buffers recom-

Table 1: Number of Amino Acids Whose p*K*s Were Used To Compute the Titration Curves in TFF1 and TFF3

		number	
amino acid	p*K*	TFF1	TFF3
arginine	12.40	3	3
lysine	10.50	1	3
tyrosine	10.00	1	2
α-NH_3	9.69	1	1
histidine	6.00	0	1
glutamic acid	4.25	7	5
aspartic acid	3.86	3	3
α-COOH	2.34	1	1
total		60	59

mended by Pharmacia for cation exchange between pH 3.5 and pH 5.0 and for anion exchange between pH 5.0 and pH 10.5. They were then loaded at 2 mL/min onto Mono S and Mono Q columns, respectively. Elution was with a 0−400 mM NaCl gradient at 1 mL/min.

RESULTS

Production of Recombinant Monomeric and Dimeric TFF3. Recombinant TFF3 was produced in *E. coli* using the pEZZ18 vector that we have used previously to express TFF1 (*17*). For the production of monomeric TFF3, dimerization is prevented by the replacement of the extra-trefoil domain Cys^{57} with the isosteric amino acid serine (Figure 1A). TFF3 is expressed as a fusion protein that is translocated to the periplasmic space where the oxidizing environment facilitates correct folding. The fusion partner comprises two "Z domains" which permit purification of the fusion protein from bacterial proteins in the periplasmic space by affinity chromatography on IgG−Sepharose. The fusion protein is cleaved by digestion with factor Xa and TFF3 purified from the fusion partner and other proteins that bind to IgG−Sepharose by a second passage over an IgG−Sepharose column. At this stage in the purification procedure recombinant TFF3 Cys^{57} and TFF3 Ser^{57} analogues migrate as single discrete bands and are >95% pure when analyzed on denaturing polyacrylamide gels (data not shown).

The homogeneity of the purified recombinant TFF3 Ser^{57} is demonstrated by its migration as a single protein band after electrophoresis under nondenaturing conditions (Figure 1B) and its elution as a discrete peak after anion-exchange chromatography (Figure 1D). In contrast, TFF3 Cys^{57} was more heterogeneous when analyzed by both nondenaturing gel electrophoresis (Figure 1B) and anion-exchange chromatography (Figure 1E). The heterogeneity of TFF3 Cys^{57} was reduced by treatment with thiol agents. Overnight incubation at 4 °C with 1 mM DTT (Figure 1C) or β-mercaptoethanol (data not shown) was very effective at converting TFF3 Cys^{57} to a form that comigrated with the Ser^{57} analogue and was more effective than cysteine or combinations of reducing and oxidizing glutathione (data not shown). The observation that TFF3 Cys^{57} migrates as a single band following treatment with a thiol agent (Figure 1C) suggested that the heterogeneity results from inappropriate disulfide bond formation. Analysis of TFF3 Cys^{57} treated with 1 mM DTT by anion-exchange chromatography showed that it was eluted as a major peak by a NaCl concentration similar to that required to elute TFF3 Ser^{57} (Figure 1F).

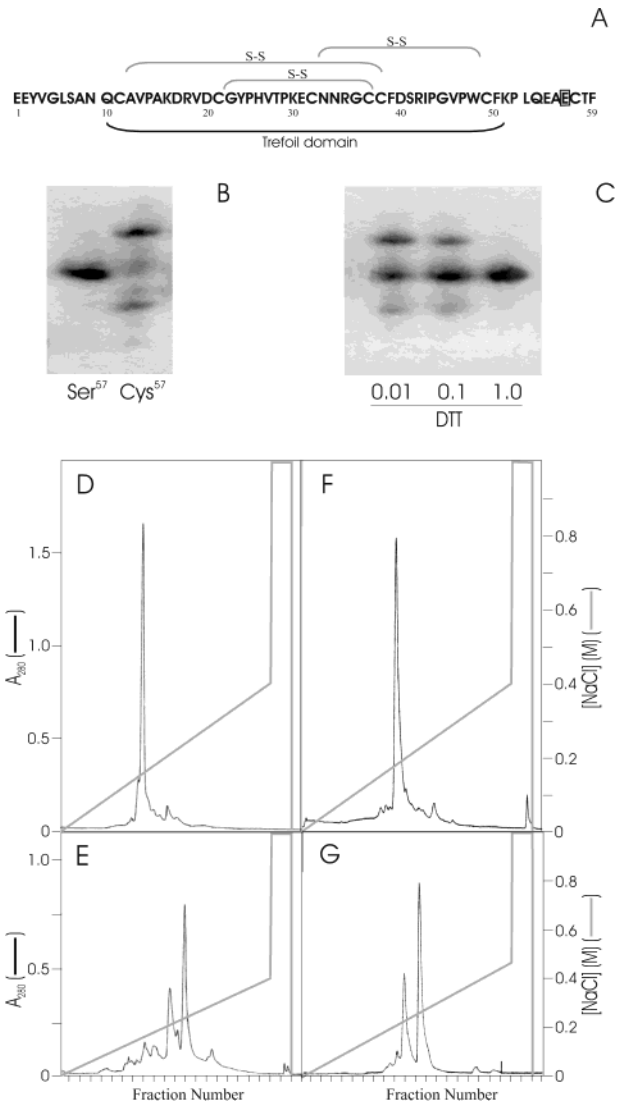

FIGURE 1: Production and purification of the TFF1 Ser^{58} and TFF3 Cys^{57} proteins. (A) Sequence of the TFF3 protein with the intramolecular disulfide bonds between the six cysteines in the trefoil domain shown by brackets (S−S). The position of the trefoil domain is indicated below the sequence, and the extra-trefoil domain cysteine that is substituted with the isosteric amino acid serine in TFF3 Ser^{57} for the production of monomeric TFF3 is highlighted with a shaded box. (B) TFF3 Ser^{57} and TFF3 Cys^{57} were electrophoresed on a nondenaturing polyacrylamide gel and stained as described in the Experimental Procedures. (C) 2 μg of TFF3 Cys^{57} was incubated overnight at 4 °C in 10 μL of 100 mM NaCl and 50 mM Tris-HCl, pH 7.8, with 0.01 mM DTT, 0.1 mM DTT, or 1 mM DTT prior to electrophoresis under nondenaturing conditions. (D) TFF3 Ser^{57} was loaded onto a 1 mL Mono Q ion-exchange chromatography column and eluted with a salt gradient as described in the Experimental Procedures. (E) TFF3 Cys^{57} was loaded onto a 1 mL Mono Q column in 20 mM Tris-HCl, pH 8.5, and eluted with a salt gradient. (F) TFF3 Cys^{57} was incubated overnight in 1 mM DTT at 4 °C prior to chromatography. (G) Fractions of the TFF3 Cys^{57} eluted protein peak were diluted 10-fold in 20 mM Tris-HCl, pH 8.5, loaded onto a 1 mL Mono Q column, and eluted with a salt gradient.

Rechromatography on anion-exchange resin of the purified peaks of recombinant TFF3 showed that the profile of the TFF3 Ser^{57} elution peak did not change (data not shown). In contrast, TFF3 Cys^{57} eluted as two peaks, one at the same salt concentration and one at a higher salt concentration (Figure 1G). Analysis of the protein in the peaks by denaturing gel electrophoresis confirmed that it had a

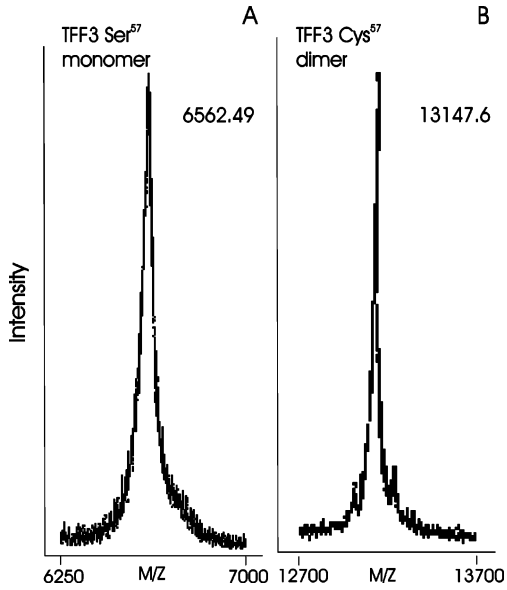

FIGURE 2: Matrix-assisted laser desorption MS of TFF3 Ser^{57} and TFF3 Cys^{57}. Spectra of the TFF3 Ser^{57} monomer (A) and TFF3 Cys^{57} dimer (B) were collected. The numbers above the peaks indicate the measured molecular masses.

Table 2: Physical Characteristics of TFF1 and TFF3 Monomers and Dimers

	TFF1		TFF3	
properties	monomer (Ser^{58})	dimer (Cys^{58})	monomer (Ser^{57})	dimer (Cys^{57})
predicted chemical mass (Da)	6652	13334	6558	13147
molecular mass by ultracentrifugation (Da)	5800 ± 300	14100 ± 700	6426 ± 300	13064 ± 600
molecular mass by MS (Da)	6645	13378	6562	13148
Stokes radius (Å)	15.5	20.8	13.7	17.8
frictional coefficient[a]	1.25	1.34	1.12	1.14

[a] The statistical significance of the differences between the frictional coefficients of the TFF1 monomer and TFF3 monomer and of the TFF1 dimer and TFF3 dimer was $p < 0.001$.

molecular mass of ~6.5 kDa and that the different elution was not explained by degradation (data not shown). Electrophoresis under nondenaturing conditions showed that the TFF3 that required a higher salt concentration to elute migrated more slowly, consistent with it having dimerized via a disulfide bond between two Cys^{57} residues.

The molecular masses of TFF3 Ser^{57} and the slower migrating form of TFF3 Cys^{57} were determined by matrix-assisted laser desorption MS (Figure 2). The measured molecular mass of TFF3 Ser^{57} was 6562.49 Da (Figure 2A), which is close to the theoretical molecular mass of 6558 Da (Table 2). The molecular mass of the higher molecular weight form of TFF3 Cys^{57} was 13147.6 Da (Figure 2B), which is close to the theoretical molecular mass of 13146.8 Da for a homodimer formed via a disulfide bond between two Cys^{57} residues (Table 2). In the experiments described below, the nondimerizing recombinant TFF3 Ser^{57} is used to characterize the physical properties of the TFF3 monomer, and the dimeric form purified away from the monomeric form of TFF3 Cys^{57} is used to characterize TFF3 dimers.

Comparison of the Hydrodynamic Properties of TFF1 and TFF3. Analytical centrifugation was used to confirm the monomeric and dimeric nature of the purified TFF3 Ser^{57} and TFF3 Cys^{57} recombinant proteins under native conditions in solution. The noninteracting model was adequate to describe all of the data. TFF3 Ser^{57} behaved as an ideal solute with an anhydrous molecular mass of 6426 ± 300 Da (Table 2). This agrees well with the predicted chemical mass of TFF3 Ser^{57}, which is 6558 Da, and is consistent with there being no aggregation of the protein at any concentration. The slower migrating form of TFF3 Cys^{57} also behaves as an ideal solute but with an anhydrous molecular mass of 13064 ± 600 Da (Table 2). The predicted chemical mass of a homodimer of TFF3 Cys^{57} formed via a disulfide bond between two Cys^{57} residues is 13147 Da. This confirms that the slower migrating form of TFF3 Cys^{57} is a dimer.

We have shown previously that recombinant TFF1 also forms homodimers via the extra-trefoil domain Cys^{58} residue in the flexible carboxy terminus and that the TFF1 Ser^{58} analogue is monomeric (*10*; Table 2). The hydrodynamic properties of the purified TFF1 recombinant proteins were compared with those of the TFF3 monomers and dimers. The molecular masses of the two monomers differ by less than 100 Da and the two homodimers by less than 200 Da (Table 2).

The Stokes radii of the TFF1 monomer and dimer and the TFF3 monomer and dimer were measured by gel filtration on Superdex 75 (Figure 3 and Table 2). Aprotinin, a globular protein that has a molecular mass similar (6500 Da) to those of TFF1 and TFF3 (Table 2) eluted at 15 mL. The TFF1 monomer eluted earlier at 13.25 mL whereas the TFF3 monomer eluted in a more similar position to aprotinin at 14.2 mL. The Stokes radii of the TFF1 and TFF3 monomers calculated from the elution volumes were 15.5 and 13.7 Å, respectively. The Stokes radii were used to calculate the frictional coefficients as described by Siegel and Monty (*18*) using the theoretical molecular masses, as these corresponded well to the values obtained by ultracentrifugation and mass spectrometry. The frictional coefficient of TFF1 was 1.25 and of TFF3 was 1.12, suggesting that neither protein is markedly asymmetric but showing that TFF1 is more asymmetric than TFF3 ($p < 0.001$). The globular protein ribonuclease A has a molecular mass similar to those of the TFF1 and TFF3 dimers and eluted at 13 mL. The TFF1 dimer eluted earlier at 11.65 mL and the TFF3 dimer at 12.6 mL, corresponding to Stokes radii of 20.8 and 17.8 Å, respectively. The frictional coefficients were 1.34 for the TFF1 dimer and 1.14 for the TFF3 dimer, showing that, as for the monomeric forms of the proteins, the TFF1 dimer is more asymmetric than the TFF3 dimer ($p < 0.001$). The difference in asymmetry between the TFF1 dimer and TFF1 monomer is larger than between the two forms of TFF3.

TFF1 and TFF3 migrate in similar positions on denaturing polyacrylamide gels, as would be expected for proteins whose molecular masses are so close (Figure 4A). Surprisingly, the TFF1 monomer migrates considerably faster than the TFF3 monomer when they are not reduced or denatured with SDS (Figure 4B). The faster migration of TFF1 suggests that there may be a larger difference in charge between the two proteins than suggested by the difference between their theoretical isoelectric points. The TFF1 dimer migrated considerably faster than the TFF3 dimer (Figure 4B).

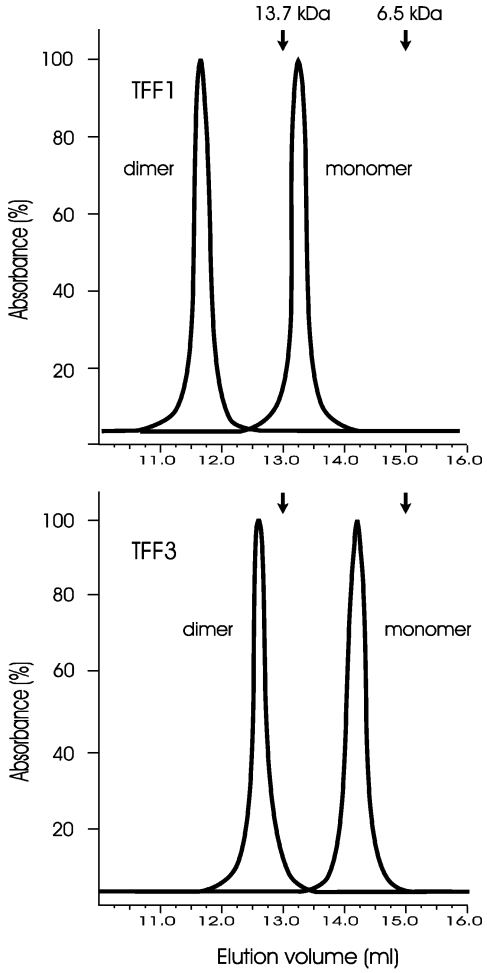

FIGURE 3: Gel filtration of TFF1 and TFF3 monomers and dimers. TFF1 and TFF3 proteins were analyzed by gel filtration on a Pharmacia Superdex 75 gel-filtration column in 50 mM sodium phosphate buffer, pH 6.5. Absorbance was measured at 280 nm, and the elution volumes are shown on the abscissa. The elution volumes of two of the globular protein standards aprotinin (6.5 kDa) and ribonuclease A (13.7 kDa) are indicated by arrows.

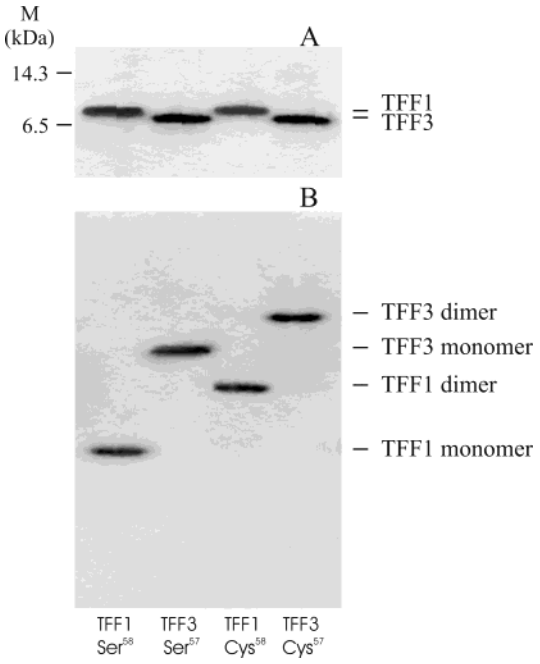

FIGURE 4: Relative migration of TFF1 and TFF3 during electrophoresis under nondenaturing conditions. Recombinant TFF1 and TFF3 monomers and TFF1 and TFF3 dimers were purified and electrophoresed on denaturing (A) or nondenaturing (B) polyacrylamide gels, and the gels were stained as described in the Experimental Procedures. The sizes of the molecular mass markers (M) are shown on the left-hand side. The positions of denatured TFF1 and TFF3 (A) and of native TFF1 and TFF3 monomers and dimers (B) are shown on the right-hand side.

Differences in the Overall Charge of TFF1 and TFF3. To investigate the possibility that the overall charge of TFF1 and TFF3 is more dissimilar than their theoretical isoelectric points suggest, titration curves were computed. There are dramatic differences between the curves computed for the two proteins (Figure 5A,B). At low pH, TFF1 is predicted to have a lower charge than TFF3 because it has fewer basic amino acids. At neutral pH, TFF1 is predicted to be much more negatively charged than TFF3 because it has a higher ratio of acidic to basic groups. At high pH they are predicted to have a similar overall charge because they contain similar numbers of positively charged amino acids. The computed titration curves account for the differences in migration of TFF1 and TFF3 (Figure 4B), as the buffers used for nondenaturing gel electrophoresis are at pH 6.8 and 8.8 where there is a large difference in the charge of the two proteins. TFF1 is three times more negatively charged than TFF3 at pH 7. This is consistent with the faster migration of TFF1, despite its larger Stokes radius, compared to TFF3.

The difference in overall charge between TFF1 and TFF3 was tested experimentally by ion-exchange chromatography of the two proteins at different pH values using Mono S and Mono Q ion-exchange resins. At pH 3.5, both proteins interact with the negatively charged Mono S resin (Table 3), but a higher salt concentration was required to elute TFF3. This suggests that TFF3 is bound more strongly than TFF1, reflecting its greater overall positive charge. At pH 4.0 the majority of TFF1 bound to the resin, but the interaction is weaker than at pH 3.5 and much weaker than that of TFF3 at pH 4.0. At pH 4.5 a significant proportion of TFF3 binds to the Mono S resin, albeit weakly, but the majority of TFF1 binds to the positively charged Mono Q resin. This interaction occurs only 0.5 pH unit above the predicted isoelectric point of TFF1 but is consistent with the precipitous drop in predicted charge of TFF1 immediately above its isoelectric point (Figure 5A). Thereafter, TFF1 binds strongly to the Mono Q resin up to pH 6.5. At pH 5.0, TFF3 does not bind to either resin, which is consistent with its theoretical isoelectric point of 4.75. TFF3 interacted weakly with the Mono Q resin as the pH was increased until pH 8.0 when it bound strongly. This shows that the overall charge of TFF3 is sufficiently low to permit a strong interaction with the resin at pH 8.0 when it is predicted to have a charge of -2. Above pH 9.0 the interaction of TFF3 with the resin increases with increasing pH, which is consistent with the predicted decrease in overall charge of the protein in this pH range (Figure 5B). Thus these empirical data for the overall charge characteristics of TFF1 and TFF3 agree well with the computed theoretical values.

The charges of the TFF1 and TFF3 dimers were also considered. Comparison of the interaction of the TFF1 dimer and monomer with Mono S at pH 3.5 shows that the dimer interacts more strongly than the monomer (Table 4). This is in agreement with the predicted overall charge for the dimer of $\sim+5$ compared to $\sim+2$ for the monomer. The interaction

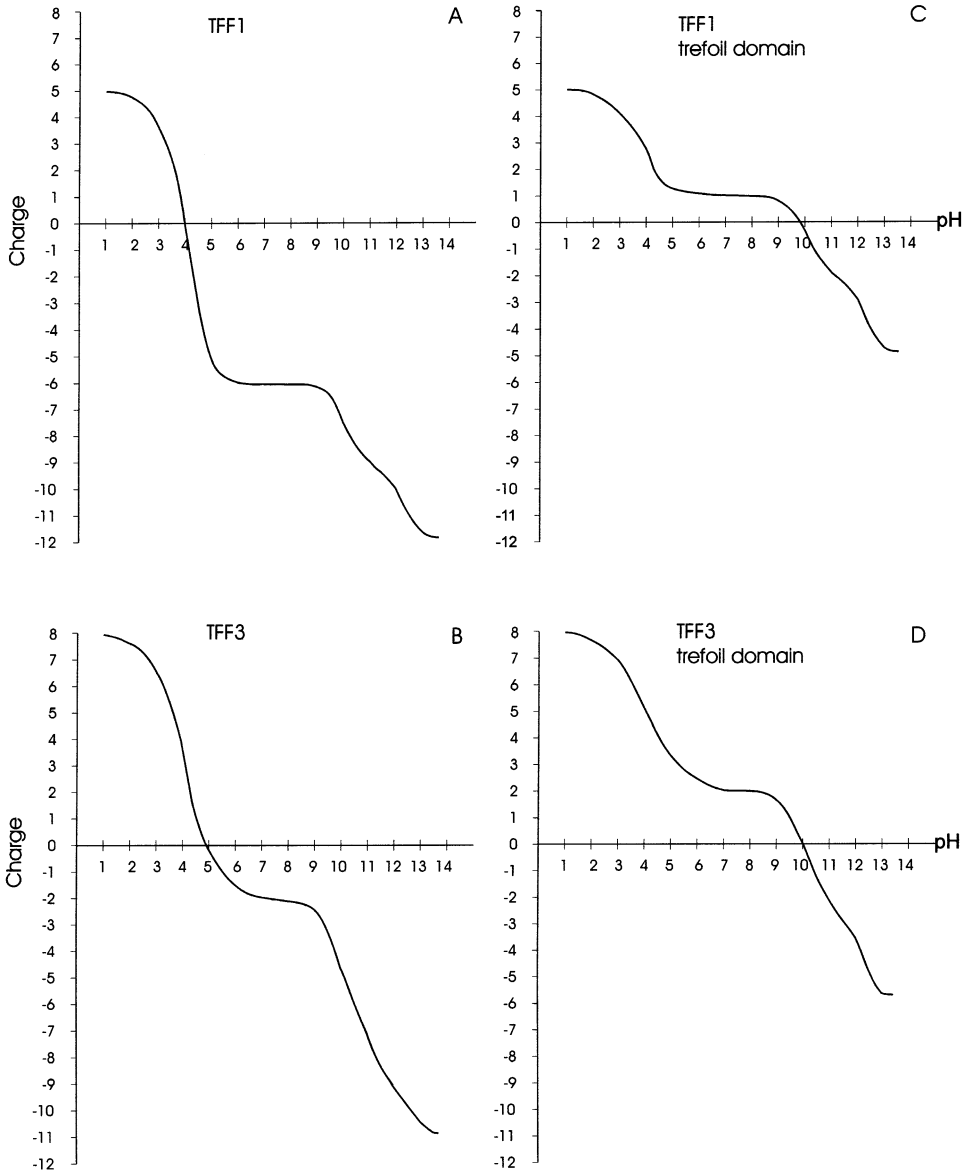

FIGURE 5: Titration curves for TFF1 and TFF3. Theoretical titration curves for TFF1 (A) and TFF3 (B) were computed using the p*I* values for the amino acids shown in Table 1. The titration curves for the trefoil domains of TFF1 and TFF3 are shown in panels C and D, respectively.

of the dimer is also stronger at pH 4.0. Above pH 5.0, the TFF1 dimer interacts very strongly with the anion-exchange resin, and NaCl concentrations above 300 mM are required to elute the protein. This is consistent with a predicted overall negative charge of −11 at pH 5.0.

Differences in Charge Distribution between TFF1 and TFF3. The charged residues in the termini of both proteins are exclusively acidic, but whereas TFF3 has an equal number of acidic residues in the termini and in the trefoil domain, TFF1 has a greater number and a higher proportion in the termini. TFF3 contains more charged residues than TFF1 in the trefoil domain, and a greater proportion are basic. Thus TFF1 has more acidic residues than TFF3, and all of the excess is outside the trefoil domain. In contrast, TFF3 has more basic residues than TFF1, and all of the excess is in the trefoil domain.

The steeper gradient of the titration curve for the TFF3 trefoil domain reflects the greater percentage of charged residues that it contains whereas the relatively flat curve for the TFF1 trefoil domain is due to the low number of charged residues. This reinforces the observation that the trefoil domain of TFF3 contributes more significantly to the overall charge characteristics of the TFF3 protein than does the trefoil domain of TFF1.

The structures of both of the TFF1 and TFF3 monomers are published. The amino and carboxy termini of both proteins are relatively unstructured, and all of their acidic residues are exposed to the solvent. In the trefoil domain, there is a conserved salt bridge formed between Arg14 and Asp35 in TFF1 and the equivalent residues, Arg18 and Asp39, in TFF3. The side chains of these conserved residues are buried and the charges neutralized by the formation of the salt bond. The remaining charged residues in the trefoil domains, two acidic and three basic in TFF1 and three acidic and six basic in TFF3 have solvent-exposed side chains.

The electrostatic potentials mapped to the surfaces of TFF1 and TFF3 are shown in Figure 6B, and comparison reveals the considerable differences between the two molecules. TFF1 is markedly polarized, with the trefoil domain being neutral or positively charged while the N- and C-termini are

Table 3: Interaction of TFF1 and TFF3 with Ion-Exchange Resins at Different pH Values

		TFF1		TFF3	
matrix	pH	% bound	[NaCl] (mM)	% bound	[NaCl] (mM)
Mono S	3.5	100	160	100	330
	4.0	>90	90	>90	220
	4.5	5	90	~80	130
	5.0	nd[a]		<5	
Mono Q	4.5	90	180	nd	
	5.0	100	240	<5	
	5.5	100	270	~5	70
	6.0	100	250	~5	90
	6.5	100	260	<5	120
	7.0			~30	100
	7.5			~80	120
	8.0			100	140
	8.5			100	140
	9.0			100	150
	9.5			100	170
	10.0			100	200
	10.5			100	250

[a] nd = not determined.

Table 4: Comparison of the Interactions of the TFF1 Monomer and Dimer with Ion-Exchange Resins at Different pH Values

		TFF1			
		monomer		dimer	
matrix	pH	% bound	[NaCl] (mM)	% bound	[NaCl] (mM)
Mono S	3.5	100	160	100	290
	4.0	>90	90	100	190
	4.5	5	90	>90	90
	5.0	nd[a]		~20	90
Mono Q	4.5	90	180	nd	
	5.0	100	240	nd	
	5.5	100	270	100	370
	6.0	100	250	100	345
	6.5	100	260	100	370

[a] nd = not determined.

strongly negatively charged. In contrast, the distribution of surface charge on TFF3 is more heterogeneous, with the basic and acidic side chains distributed more evenly over the surface of the trefoil domain.

Finally, we examined the locations and orientations of the side chains of individual charged residues in the trefoil domains of the two molecules. Of the conserved or partially conserved residues, Arg12 in TFF1 and Lys16 in TFF3 are located in the same positions in loop 1, but their side chains are orientated differently (Figure 6C and Table 6). The positions of Glu13 and Asp17 in loop 1 of TFF1 and TFF3, respectively, are the same, and their side chains are both bent back toward the termini. Arg14 and Arg18 are the structurally important amino acids that form salt bridges with Asp35 and Asp39 in TFF1 and TFF3, respectively. Lys30 in TFF1 and Arg34 in TFF3 are conserved basic residues positioned in the short α-helix of loop 2, and the side chains are similarly orientated. Arg39 in TFF1 and Arg41 in TFF3 do not have the identical location in the primary sequence but are located in similar positions at the tip of loop 3 and could conceivably perform a similar function in the two proteins. But while the side chain of Arg39 points directly out from the front of TFF1 away from the termini, that of Arg41 in TFF3 is turned through 180° to lie along the surface of the trefoil domain and points toward the termini. Thus among these four nonstructural residues, the side chains of two adopt similar orientations and two very different orientations in TFF1 and TFF3.

The nonconserved Asp36 in TFF1 is located between loops 2 and 3 (Figure 6D). Among the charged residues unique to TFF3, Asp20 is between loops 1 and 2. His25 is located at the tip of loop 2, and the side chain folds back toward the α-helix. Lys29 and Glu30 along with the conserved Arg34 are in the short α-helix in loop 2, and all of the side chains point in the same direction away from the trefoil domain to give a local concentration of charged groups. Finally, Lys50 lies at the end of loop 3, and its side chain points into the solvent.

DISCUSSION

TFF1 and TFF3 are closely related proteins of almost identical size which, although expressed in different places, have similar biological activities in various model systems both in vitro and in vivo. TFF1-null mice are characterized by an absence of gastric antral and pyloric mucus and have morphological abnormalities in the antrum leading to formation of adenomas and carcinomas (*15*), TFF3-null mice have a reduced ability to repair damage to the intestinal mucosa (*16*). This indicates that the physiological functions of TFF1 and TFF3 may be quite different and that they may elicit their actions through distinct protein−protein interactions. In the present study we have investigated whether there are differences in the hydrodynamic and charge properties of the two molecules that may reflect their different biological profiles.

Production of TFF3. We have described the preparation of the pure recombinant TFF3 monomer and dimer using an *E. coli* expression system. Two other groups have described the production of recombinant human TFF3. Thim et al. (*11*) produced recombinant TFF3 using yeast, and Kinoshita et al. (*19*) produced TFF3 fusion proteins with a 14 kDa TRX domain at the N-terminus in *E. coli*. The pEZZ18 system utilized in this study has been used to produce protein having the correct N-terminus. The molecular masses of the TFF3 monomer and dimer measured by matrix-assisted laser desorption MS were very close to the theoretical values, and the traces show less heterogeneity than those published previously (*11*).

The two proteins will be valuable for characterizing the functional roles of TFF3 and for elucidating the signal transduction pathways of trefoil proteins. The proteins have also been useful for structural studies, and the pEZZ18 vector makes this an ideal system with which to perform structure−function analyses because it allows the production of single-stranded DNA for in vitro mutagenesis.

Hydrodynamic Properties. As the size and shape of the trefoil domains are almost identical, the differences between the Stokes radii and frictional coefficients of TFF1 and TFF3 monomers may be ascribed to the longer carboxy terminus of TFF1.

The amino and carboxy termini are juxtaposed in the NMR structures of both proteins; however, the TFF1 molecule is longer than the TFF3 molecule because the carboxy-terminal

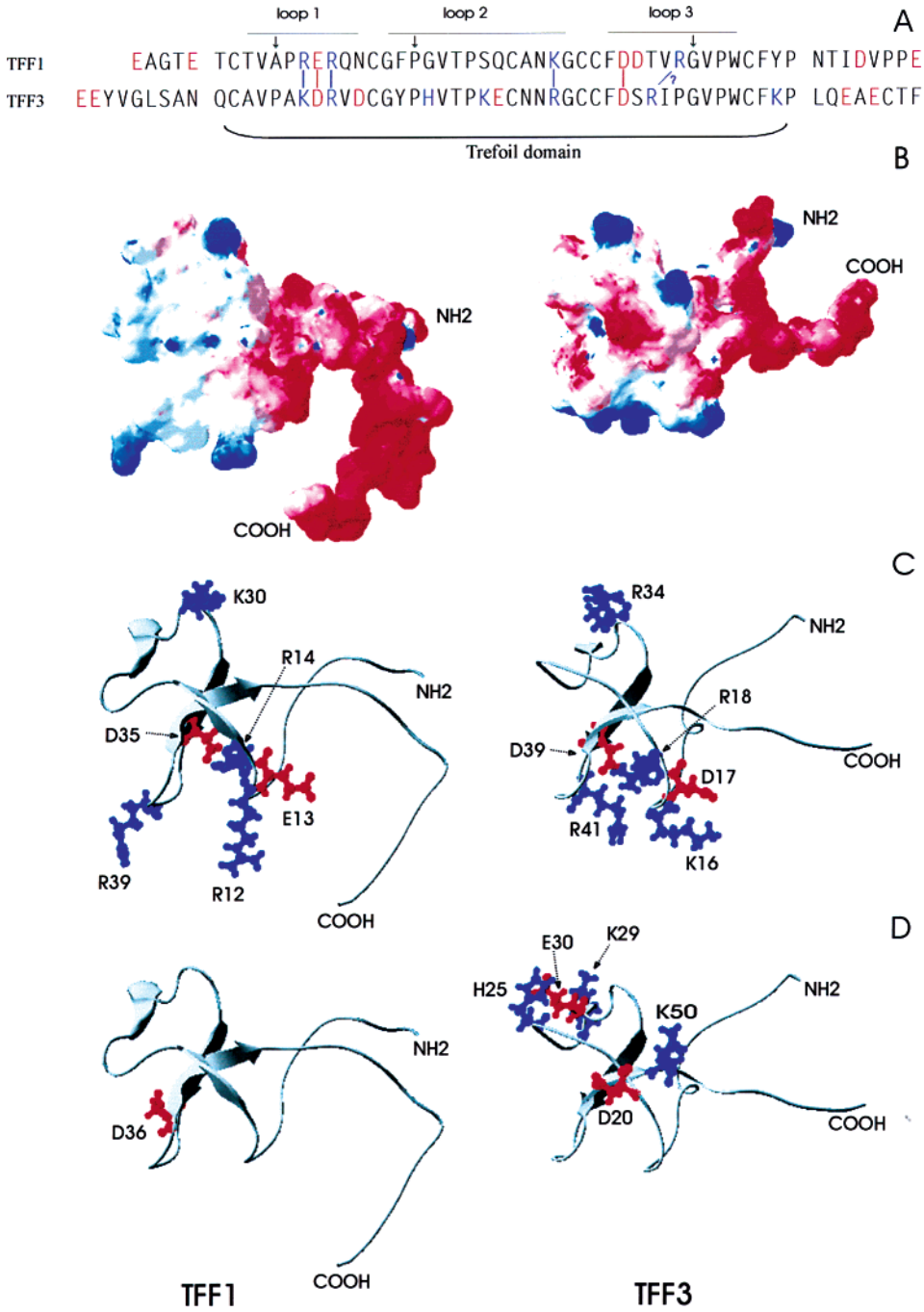

FIGURE 6: Distribution of charged amino acids in TFF1 and TFF3. (A) Sequences of the TFF1 and TFF3 proteins with the trefoil domains and structural loops 1, 2, and 3 demarcated. The charged residues are colored (E and D, red; R, K, and H, blue). Vertical lines connect conserved charged residues within the trefoil domain. The approximate extents of the three loops are indicated above the sequences, and arrows indicate the amino acids at the tips of the loops. (B) Electrostatic potential mapped to the surface of TFF1 and TFF3. Significant negative charge is shown in red, significant positive charge in blue, and neutral as white. (C) The location and orientation of the conserved charged residues in TFF1 and TFF3 are indicated. Amino acids with acidic side chains are shown in red, and those with basic side chains are in blue. (D) The location and orientation of the nonconserved charged residues are indicated.

Table 5: Distribution of Residues with Charged Groups in TFF1 and TFF3

	N-terminus		trefoil domain		C-terminus	
residues	TFF1	TFF3	TFF1	TFF3	TFF1	TFF3
acidic	2	2	3	4	5	2
basic	0	0	4	7	0	0
total no.	5	9	42	42	13	8
% charged	40	22	17	26	38	25

domain of TFF1 is four and six amino acids longer than the amino- and carboxy-terminal domains, respectively, of TFF3 (Figure 6B). In addition, the carboxy terminus of TFF1 is very positively charged, which may be predicted to encourage an extended conformation and increase the asymmetry of the molecule. Although structural and hydrodynamic data are difficult to reconcile, estimates of the maximum diameters from the NMR structures, measurements of the distance between Phe60 and Phe19 for TFF1 and between Phe59 and Pro24 for TFF3, do suggest that the diameter of TFF1 is approximately 20% greater than that of TFF3, which is consistent with the hydrodynamic measurements.

Table 6: Conservation of Individual Residues with Charged Groups in TFF1 and TFF3

domain	charged residue TFF1	charged residue TFF3	surface	conservation of amino acids within trefoil domain	orientation of conserved amino acid side chains in trefoil domain
N-terminus	Glu1	Glu1	yes	not applicable	not applicable
	Glu5	Glu2	yes	not applicable	not applicable
trefoil domain	Arg12	Lys16	yes	conserved	different
	Glu13	Asp17	yes	conserved	similar
	Arg14	Arg18	no	conserved	buried
		Asp20	yes	unique	not applicable
		His25	yes	unique	not applicable
		Lys29	yes	unique	not applicable
		Glu30	yes	unique	not applicable
	Lys30	Arg34	yes	conserved	similar
	Asp35	Asp39	no	conserved	buried
	Asp36		yes	unique	not applicable
	Arg39	Arg41	yes	conserved?	opposite
		Lys50	yes	unique	not applicable
C-terminus	Asp51		yes	not applicable	not applicable
	Glu55	Glu54	yes	not applicable	not applicable
	Glu56		yes	not applicable	not applicable
	Glu57	Glu57	yes	not applicable	not applicable
	Glu59		yes	not applicable	not applicable

The TFF1 dimer is significantly more asymmetric than the TFF3 dimer; the frictional ratios are 1.33 and 1.14, respectively ($p < 0.001$). Indeed, the difference in symmetry is more marked for the dimers than for the monomers. The longer carboxy terminus of TFF1 may also explain the increase in asymmetry of the TFF1 dimer compared to the TFF3 dimer. If the two trefoil domains are separated by their extended carboxy termini, then the shorter carboxy terminus in TFF3 would cause the two trefoil domains to be in closer proximity.

The structure of the TFF1 dimer has been published recently (*13*) and shows that the two trefoil domains are at opposite ends of a flexible linker. The observation that the TFF3 dimer is less asymmetric and has a markedly smaller Stokes radius than the TFF1 dimer suggests that the TFF3 dimer forms a more compact structure than that of the TFF1 dimer, and it will be interesting to determine whether it more closely resembles the TFF1 dimer or TFF2.

The TFF1 dimer interacts with human gastric mucus (*23*). There is no published evidence for an interaction between TFF3 and mucus. It is possible that greater separation of the trefoil domains in the TFF1 dimer facilitates this interaction. The different frictional coefficients of the two proteins would also be consistent with their interaction with distinct binding proteins and receptors.

Differences in Charge. There are marked differences in the overall charges of TFF1 and TFF3. The experimental data obtained over a wide pH range agree well with the computed p*I* values for both molecules. The significant difference in the charge of the two proteins necessitated the development of different purification protocols for recombinant TFF1 and TFF3 and accounts for the large differences in their mobility in nondenaturing polyacrylamide gels. The residues outside the trefoil domains of both TFF1 and TFF3 are acidic and contain no basic amino acids. This is in marked contrast to the trefoil domains which, particularly in TFF3, contain an excess of basic residues. In addition, the different orientations of the side chains of the charged amino acids exacerbate the greater polarization of TFF1. The relevance of these charged residues to the biological activities of TFF1 and TFF3 is difficult to assess as, apart from the extra-trefoil domain cysteine, the surfaces and residues of the trefoil domains involved in intermolecular interactions are not known.

In TFF1, Asp36 (Figure 6D) contributes to an area of negative charge on the surface of the molecule which is largely absent on TFF3. Arg39 is located on the end of loop 3 and creates a positively charged protrusion into the solvent adjacent to the conserved hydrophobic patch comprising Phe19, Pro20, Pro42, and Trp43. These residues may contribute significant differences to the surface topology and potential of the two proteins. In TFF3, Lys29 and Glu30 together with the conserved Arg34 are all located in the short six-residue α-helix which extends from residue 29 to residue 34.

Interest has focused previously on the hydrophobic residues including Tyr23, Pro24, Val45, Pro46, Trp47, and Phe49. They form a cluster of conserved residues on the surface in the region of the ends of loops 2 and 3, and it has been suggested that they are involved in protein−protein interactions. Interestingly, the conformation of the tip of loop 2 is significantly different in the three human trefoil proteins, which may contribute to the specific biological effects of individual trefoil proteins. It is noteworthy that all of the charged residues in this region in TFF1 and TFF3 are unique. It could be envisaged that these residues might be involved in individualizing the protein−protein interactions and biological activities of the trefoil family of proteins.

Physiological Relevance. The data described in this study show that both the monomeric and dimeric forms of TFF1 have very different charge properties from the corresponding forms of TFF3. It is possible that these differences between TFF1 and TFF3 reflect adaptation to the different environments in which they function. TFF3 is expressed predominantly in the goblet cells of the small and large intestine and would be predicted to be exposed to pHs of between 6 and 8 (*20*, *21*). TFF1 is synthesized by the mucus secretory cells of the gastric epithelium (*22*, *23*) and is found in the adherent mucous gel layer (*24*) and in the gastric lumen (*22*). TFF1 is therefore exposed to pHs between 7 and 1 at various stages in the transition from inside the cells across the pH gradient that is maintained in the adherent mucus layer and into the gastric lumen. During this transition there will be a large shift in the overall charge of the protein, particularly outside the trefoil domain (Figure 5A,C). Decrease in pH has been shown to promote folding of acidic domains (*25*−*27*), due to protonation of acidic side chains. It is possible that a similar process occurs in the carboxy terminus of TFF1. This could be of functional significance during exposure of TFF1-bound secreted mucins to low pH following the release of secretory granules.

In conclusion, we have shown experimentally that TFF1 and TFF3 have markedly different hydrodynamic properties and differences in the overall and topological location of their surface charge. The experimental observations agree with known structural features of the two proteins. This strengthens the supposition that differences in conformational and charge characteristics of the two molecules account for their different biological activities.

ACKNOWLEDGMENT

We thank Dr. Andrew Lane (National Institute for Medical Research, Mill Hill, London, U.K.) for assistance with and discussion of the ultracentrifugation experiments.

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*, 7895−7903.
2. May, F. E. B., and Westley, B. R. (1986) Cloning of estrogen regulated mRNA sequences from human breast cancer cells, *Cancer Res. 46*, 6034−6038.
3. May, F. E. B., and Westley, B. R. (1997) Expression of human intestinal trefoil factor in malignant cells and its regulation by oestrogen in breast cancer cells, *J. Pathol. 182*, 404−413.
4. Gruvberger, S., Ringner, M., Chen, Y., Panavally, S., Saal, L. H., Borg, A., Ferno, M., Peterson, C., and Meltzer, P. S. (2001) Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns, *Cancer Res. 61*, 5979−5984.
5. West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., Zuzan, H., Olson, J. A., Jr., Marks, J. R., and Nevins, J. R. (2001) Predicting the clinical status of human breast cancer by using gene expression profiles, *Proc. Natl. Acad. Sci. U.S.A. 98*, 11462−11467.
6. 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.
7. May, F. E. B., and Westley, B. R. (1997) Trefoil proteins: their role in normal and malignant cells, *J. Pathol. 183*, 4−7.
8. Poulsom, R., and Wright, N. A. (1993) Trefoil peptides: a newly recognized family of epithelial mucin-associated molecules, *Am. J. Physiol. 265*, G205−G213.
9. 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.
10. 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 58, *Biochem. J. 327*, 117−123.
11. Thim, L., Wöldike, H. F., Nielsen, P. F., Christensen, M., Lynch-Devaney, K., and Podolsky, D. K. (1995) Characterization of human and rat intestinal trefoil factor produced in yeast, *Biochemistry 34*, 4757−4764.
12. 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 solution structure of human pNR-2/pS2: a single trefoil motif protein, *J. Mol. Biol. 267*, 418−432.
13. 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.
14. Lemercinier, X., Muskett, F. W., Cheeseman, B., McIntosh, P. B., Thim, L., and Carr, M. D. (2001) High-resolution solution structure of human intestinal trefoil factor and functional insights from detailed structural comparisons with the other members of the trefoil family of mammalian cell motility factors, *Biochemistry 40*, 9552−9559.
15. 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.
16. Mashimo, H., Wu, D.-C., Podolsky, D. K., and Fishman, M. C. (1996) Impaired defence of intestinal mucosa in mice lacking intestinal trefoil factor, *Science 274*, 262−265.
17. 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 Ser^{58}, *Biochem. J. 308*, 1001−1007.
18. Siegel, L. M., and Monty, K. J. (1966) Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases, *Biochim. Biophys. Acta 112*, 346−362.
19. Kinoshita, K., Taupin, D. R., Itoh, H., and Podolsky, D. K. (2000) Distinct pathways of cell migration and antiapoptotic response to epithelial injury: structure−function analysis of human intestinal trefoil factor, *Mol. Cell. Biol. 20*, 4680−4690.
20. Podolsky, D. K., Lynch-Devaney, K., Stow, J. L., Oates, P., Murgue, B., DeBeaumont, M., Sands, B. E., and Mahida, Y. R. (1993) Identification of human intestinal trefoil factor. Goblet cell-specific expression of a peptide targeted for apical secretion, *J. Biol. Chem. 268*, 6694−6702.
21. Hauser, F., Poulsom, R., Chinery, R., Rogers, L. A., Hanby, A. M., Wright, N. A., and Hoffmann, W. (1993) hP1.B, a human P-domain peptide homologous with rat intestinal trefoil factor, is expressed also in the ulcer-associated cell lineage and the uterus, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6961−6965.
22. Rio, M.-C., Bellocq, J. P., Daniel, J. Y., Tomasetto, C., Lathe, R., Chenard, M. P., Batzen Schlager, A., and Chambon, P. (1988) Breast cancer associated pS2 protein: synthesis and secretion by normal stomach mucosa, *Science 241*, 705−708.
23. 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.
24. 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.
25. Dutta, K., Alexandrov, A., Huang, H., and Pascal, S. M. (2001) pH-induced folding of an apoptotic coiled coil, *Protein Sci. 10*, 2531−2540.
26. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the proteionstructure−function paradigm, *J. Mol. Biol. 293*, 321−331.
27. Uverski, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions?, *Proteins 41*, 415−427.

BI030025L