Characterization of Human and Rat Intestinal Trefoil Factor Produced in Yeast†

Lars Thim,*,‡ Helle F. Wöldike,‡ Per F. Nielsen,‡ Mogens Christensen,‡ Kathryn Lynch-Devaney,§ and Daniel K. Podolsky§

Bioscience, Novo Nordisk A/S, DK-2880 Bagsvaerd, Denmark, and Gastrointestinal Unit, Massachusetts General Hospital, Boston, Massachusetts 02114

Received October 11, 1994⊗

ABSTRACT: Intestinal trefoil factor (ITF) from human (hITF) and rat (rITF) have been produced in *Saccharomyces cerevisiae.* The DNA encoding the two peptides were cloned by polymerase chain reactions (PCR) from a human normal colon library and a rat small intestinal epithelial cell library. Recombinant plasmids were constructed to encode a fusion protein consisting of a hybrid leader sequence and the rat and human ITF sequences, respectively. The leader sequence used serves to direct the fusion protein into the secretory (and processing) pathway of the cell. The secreted recombinant hITF was found in a monomer and a dimer form, whereas the rITF was only secreted as a dimer. The secreted peptides were purified by a combination of ionic exchange chromatography and preparative HPLC. From 8 L of yeast fermentation broth, 256 mg of hITF (monomer) and 133 mg of hITF (dimer) were isolated, and from 8.7 L of fermentation broth, 236 mg of rITF (dimer) was isolated. The structure of hITF (monomer), hITF (dimer), and rITF (dimer) was determined by amino acid analyses, peptide mapping, sequence analyses, and electrospray mass spectrometry analyses. In hITF (monomer) six of the seven cysteines are disulfide-linked to form 3 disulfide bridges. Mass analysis indicated that the last cysteine residue (Cys-57) did not exist as free (-SH) cysteine, but have reacted with cysteine to form an S−S linked cystine. Sequence and mass spectrometry analyses as well as peptide mapping showed that the dimer form of both hITF and rITF is mediated by a disulfide bridge between Cys-57 residues of two monomers.

Intestinal trefoil factor, ITF[1] (Figure 1), belongs to a growing family of gastrointestinal peptides containing a characteristic domain in which six half-cystine residues in a peptide chain of 38 or 39 amino acids and residues form three disulfide bonds to create a characteristic three-leaved structure giving the peptide family its name (Thim, 1989).

The mammalian trefoil peptides known today contain either one or two trefoil domains [for reviews, see Thim (1994), Poulsom and Wright (1993), and Hoffmann and Hauser (1993)], whereas from the frog, *Xenopus laevis* peptides and proteins containing one (Hauser & Hoffman, 1991), two (Hauser et al., 1992), four (Hoffmann, 1988), or six (Hauser & Hoffmann, 1992) trefoil domains have been described. The mammalian trefoil peptides containing one domain are the breast cancer associated pS2 peptide so far known from human (Jakowlev et al., 1984, Prud'homme et al., 1985) and mouse (Lefebvre et al., 1993) and the intestinal trefoil factor so far known from human (Podolsky et al., 1993; Hauser er al., 1993) and rat (Suemori et al., 1991; Chinery et al., 1992). Spasmolytic polypeptide (SP), which contains two trefoil domains, has been described from man (Tomasetto et al., 1990), pig (Thim et al., 1982), and mouse (Tomasetto et al., 1990). In humans, the three trefoil peptides hpS2, hITF, and hSP are all expressed under normal conditions in the gastrointestinal tract: hSP and hpS2 in the epithelial mucosal layer of the stomach (Tomasetto et al., 1990; Rio et al., 1988) and hITF in the epithelial mucosal layer of the small intestine and colon (Podolsky et al., 1993).

The physiological function of the trefoil peptides is not very well understood. Increased expression of trefoil peptides in the gastrointestinal tract has been reported in several conditions involving mucosal injury such as inflammatory bowel disease (Rio et al., 1991; Poulsom et al., 1992; Wright et al., 1993) and ulceration in the stomach and duodenum (Rio et al., 1991; Hanby et al., 1993; Wright et al., 1990). Consequently, a mucosal repair function of the trefoil peptides has been suggested (Wright et al., 1993). Evidence of trefoil peptides promoting mucosal epithelial restitution after injury has recently been given (Dignass et al., 1994; Playford et al., 1995; Babyatsky et al., 1994). The mechanism by which trefoil peptides promote their repair function may be to cross-link mucin glycoproteins to form a visco-elastic gel layer resistant to digestive enzymes (Thim, 1994; Gajhede et al., 1993). We have previously produced relatively large amounts of the two domain hSP in a yeast system (Thim et al., 1993), and in the present study we have applied the same expression system for producing rat and human ITF.

EXPERIMENTAL PROCEDURES

Cloning of Rat ITF (rITF) and Human ITF (hITF). Rat ITF cDNA was cloned from a cDNA library prepared from isolated rat small intestinal cells in the λZAPII vector (Stratagene) as previously described (Suemori et al., 1991). The human ITF cDNA was cloned from a library prepared with normal human colon cDNA in the λgt11 vector

† Parts of this work were supported by grants from the National Institutes of Health (DK 46906 and DK 43351).

* To whom correspondence should be addressed.

‡ Novo Nordisk A/S.

§ Massachusetts General Hospital.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1995.

[1] Abbreviations: ITF, intestinal trefoil factor; hITF, human ITF; rITF, rat ITF; PCR, polymerase chain reactions; PSP, porcine spasmolytic polypeptide; hSP, human spasmolytic polypeptide; hpS2, human breast cancer associated polypeptide; pyrGlu, pyrrolidone carboxylic acid.

0006-2960/95/0434-4757$09.00/0

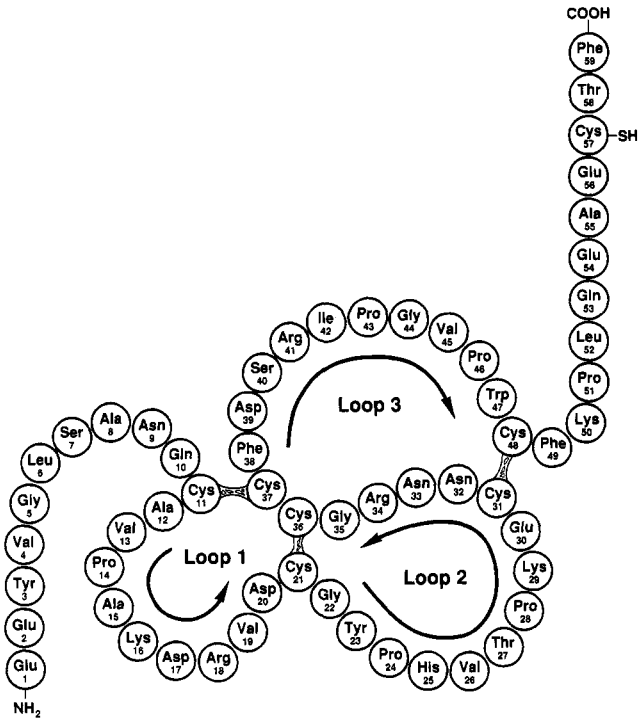

FIGURE 1: Proposed structure of human intestinal trefoil factor ITF. The primary amino acid sequence is taken from Podolsky et al. (1993) and Hauser et al. (1993). The disulfide bonds are placed in homology to PSP and pS2 (Thim, 1989).

(Clontech) screened with the full-length rITF cDNA as previously reported (Podolsky et al., 1993). Plaque-purified clones were amplified by PCR using primers prepared from the sequences flanking the λgt11 insertion site (35 cycles of 95 °C for 1 min, 48 °C for 1.5 min, and 78 °C for 2 min, followed by 72 °C for 8 min). The amplified insert was subcloned into the PCR 1000 vector (Invitrogen).

Construction of rITF and hITF Secreting Yeast Strains. The yeast expression plasmids pHW 756 containing the rITF gene and pHW 1066 containing the hITF gene were composed essentially as described for the hSP expression vector (Thim et al., 1993). The TPI promoter and terminator sequences are from *S. cerevisiae* triosephosphate isomerase gene, and the selection marker is POT from *Saccharomyces pombe* triosephosphate isomerase gene. The 53 amino acid long signal-leader sequence is as described:

MKAVFLVI.SLIGFCWAQPVTGDESSVEI
PEESLIIAENTTLANVAMAERLEKR

Immediately following this sequence is the hITF sequence, in accordance with Hauser et al. (1993) as outlined in Figure 1, or the rITF sequence:

QEFVGLSPSQCMVPANVRVDCGYPTVTSE
QCNNRGCCFDSSIPNVPWCFKPLQETECTF

in accordance with Chinery et al. (1992).

The plasmids were transformed into the yeast strain MT-663, carrying a deletion in the TPI gene, and selected for growth on glucose.

Fermentations. The transformants described above were cultivated at 30 °C for 72 h in yeast peptone dextrose (YPD) medium (Sherman et al., 1981) supplied with additional yeast extract (60 g/L). OD values at 660 nm of 153 and 232 for HW756 (rITF) and HW1066 (hITF), respectively, were reached at the end of the fermentations. The pH was adjusted to 2.5 with 1 M phosphoric acid at the end of the fermentation, and the yeast cells were removed by centrifugation at 3000 rpm for 15 min.

Purification of Recombinant rITF. The concentration of rITF in the yeast fermentation broth and fractions obtained during the purification was measured by analytical HPLC. Aliquots (usually 50−200 μL) were injected onto a Vydac 214TP54 reversed-phase C4 HPLC column (4.6 × 250 mm) equilibrated at 30 °C at a flow rate of 1.5 mL/min with 0.1% (v/v) TFA in 15% (v/v) acetonitrile. After 10 min of isocratic elution, the concentration of acetonitrile in the eluting solvent was raised to 55% over 40 min. Absorbance was measured at 214 nm. Three peaks eluting at 26.5, 27.3, and 28.2 min (Figure 2) were found to represent dimer forms of rITF (see Results). The peptides were quantified using a calibrated hSP standard (Thim et al., 1993).

From a 10-L fermenter, 8.7 L of fermentation broth was isolated by centrifugation. The supernatant was diluted with 14.8 L of distilled water to lower the conductivity. The sample was pumped onto a Fast Flow S-Sepharose (Pharmacia) column (5 × 42 cm) with a flow rate of 600 mL/h. Prior to the application, the column was equilibrated in 50 mM formic acid buffer, pH 3.7. Rat ITF was eluted from the column by 50 mM formic acid, pH 3.7, containing 50 mM NaCl. Fractions of 100 mL were collected at a flow rate of 600 mL/h and analyzed for the content of rITF. Fractions from the previous step containing rITF were pooled (2.3 L) and pumped onto an Amberchrome G-71 column (5 × 10 cm). Prior to the application, the column was equilibrated in 10 mM ammonium acetate buffer, pH 4.8, at a flow rate of 0.5 L/h. After application, the column was washed with 0.5 L of equilibration buffer and eluted with 10 mM ammonium acetate buffer, pH 4.8, containing 60% (v/v) of ethanol at a flow rate of 0.1 L/h. Fractions of 10 mL were collected and pooled according to their content of rITF. The ethanol concentration in the pool was increased from 60% (v/v) to 87% (v/v) by the addition of 2 vols of ethanol (99.9%, v/v), and rITF was precipitated by cooling the resulting mixture to −25 °C for 16 h.

The precipitate was collected by centrifugation for 1 h at 10000*g* at −25 °C and redissolved at room temperature in 130 mL of 20 mM formic acid, pH 3.0. The sample was pumped onto a Fast Flow SP-Sepharose (Pharmacia) column (5 × 20 cm) with a flow rate of 50 mL/h. Prior to the application, the column was equilibrated with 20 mM formic acid, pH 3.0. Peptides were eluted from the column by a linear gradient between 1.5 L of 50 mM formic acid, pH 3.0, and 1.5 L of formic acid, pH 3.0, containing 0.5 M NaCl. Fractions (10 mL) were collected at a flow rate of 80 mL/h, and the absorbance was measured at 280 nm. Fractions were assayed for the content of rITF. Fractions corresponding to rITF were pooled. Rat ITF was further purified by preparative HPLC. Pooled fractions (900 mL) were pumped onto a Vydac 214TP1022 C4 preparative HPLC column (2.2 × 25 cm) equilibrated in 0.1% (v/v) TFA. The peptides were eluted at 25 °C and at a flow rate of 5 mL/min with a linear gradient (540 mL) formed from MeCN/H_2O/TFA (10:89.9:0.1, v/v) and MeCN/H_2O/TFA (65:34.9:0.1, v/v). UV absorption was monitored at 280 nm, and fractions corresponding to 10 mL were collected and analyzed for the content of rITF. Fractions containing rITF were pooled, and the volume was reduced to 30% by vacuum centrifugation.

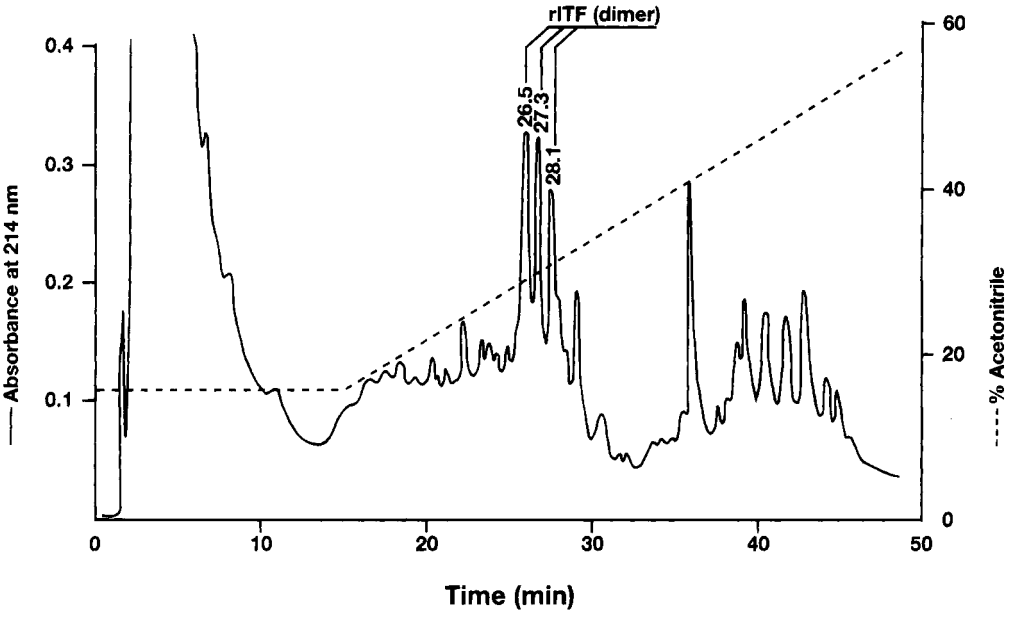

FIGURE 2: Reversed-phase HPLC on a Vydac 214TP54 column of supernatant from yeast strain HW756 expressing rat ITF.

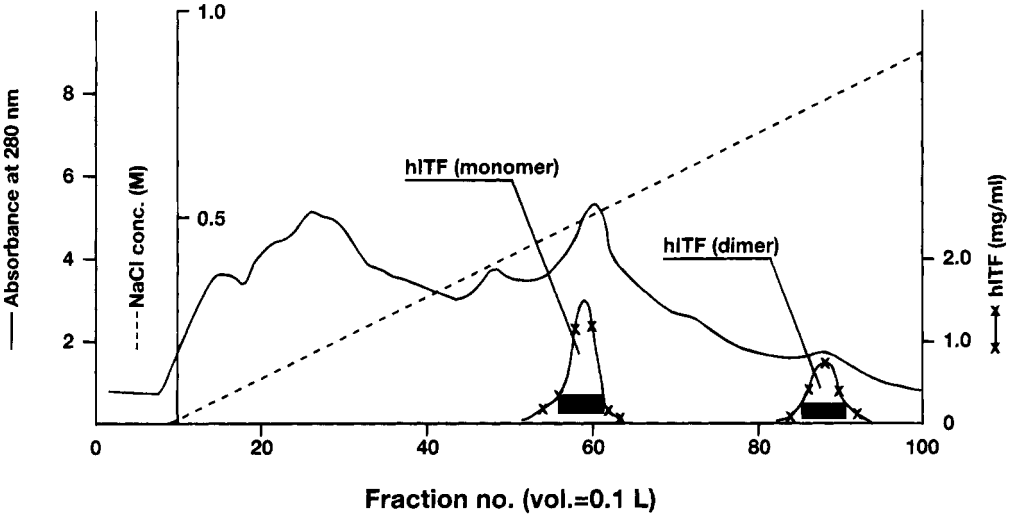

FIGURE 3: Ion-exchange chromatography on a Fast Flow SP-Sepharose column of partially purified human ITF. The amounts of hITF (monomer) and hITF (dimer) were determined by analytical HPLC. The bars indicate the fractions pooled for further purification of the monomer and dimer forms. The dashed line shows the concentration of NaCl in the eluting solvent.

From the resulting pool, rITF was isolated by lyophilization. The total yield of rITF from 8.7 L of fermentation medium was 236 mg, corresponding to an overall purification yield of 24%.

Purification of Recombinant hITF. The concentration of hITF in the yeast fermentation broth and fraction, obtained during the purification, was measured in an HPLC system identical to the one described for rITF. In this system, two peaks eluting at 21.2 and 27.1 min were found by mass spectrometry and sequence analysis to represent a dimer and a monomer form of hITF (see Results).

From a 10-L fermenter, 8.0 L of fermentation broth was isolated by centrifugation. The sample was dialyzed three times (each time in 24 h) against 40 L of 10 mM formic acid, pH 2.5. The sample was pumped (0.25 L/h) onto an Fast Flow SP-Sepharose (Pharmacia) column (5 × 40 cm). The column was washed with 5 L of 20 mM formic acid, pH 2.5, and eluted with a linear gradient formed by 5 L of 20 mM formic acid, pH 2.5, and 5 L of formic acid, pH 2.5, containing 1 M of NaCl. Fractions of 100 mL were collected and analyzed for the content of hITF (Figure 3). Two forms of hITF were eluted from the column: one representing a monomer form of hITF (eluting at 0.5 M of NaCl) and one representing a dimer form of hITF (eluting at 0.78 M of NaCl). Fractions corresponding to the two forms were pooled separately.

Each fraction was divided into three equal parts (volume, approximately 700 mL) and pumped onto a Vydac 214TP1022 C4 column (2.2 × 25 cm) equilibrated in 0.1% (v/v) TFA. The peptides were eluted at a flow rate of 4 mL/min with a linear gradient (540 mL) between MeCN/H_2O/TFA (10:89.9:0.1, v/v) and MeCN/H_2O/TFA (65:34.9:0.1, v/v). UV absorption was monitored at 280 nm, and fractions corresponding to 10 mL were collected and analyzed for the content of hITF.

Fractions from the previous step containing hITF (monomer) and hITF (dimer) were pooled separately, and the pH was adjusted to 3.0. The samples were applied separately onto an SP-Sepharose HiLoad 16/10 (Pharmacia) column (1.6 × 10 cm) equilibrated in 20 mM formic acid, pH 3.0,

containing 40% (v/v) ethanol. The column was washed with 80 mL of equilibration buffer and eluted at a flow rate of 4 mL/min with a linear gradient between 200 mL of 20 mM formic acid, pH 3.0, 40% (v/v) ethanol, and 200 mL of 20 mM formic acid, pH 3.0, 40% (v/v) containing 1 M NaCl. Fractions (5 mL) were collected and analyzed for the content of hITF.

Fractions containing hITF (monomer) and hITF (dimer), respectively, were pooled, and the peptide content was precipitated by adjusting the ethanol concentration to 90% (v/v) and cooling the mixture for 72 h at −25 °C. The precipitate was collected by centrifugation and lyophilized. The total yield from 8 L of fermentation broth was 256 mg of hITF (monomer) and 133 mg of hITF (dimer) corresponding to an overall purification yield of 50% and 65%, respectively, for the monomer and dimer forms.

Characterization of Recombinant rITF and hITF. After hydrolysis in 6 M HCl at 110 °C in vacuum-sealed tubes for 24, 48, and 96 h, peptide samples (50 μg) were analyzed on a Beckman (Model 121 MB) automatic amino acid analyzer. Half-cystine was determined as the *S*-β-(4-pyridylethyl) derivative, after reduction of the disulfide bonds by tributylphosphine (Rüegg & Rudinger, 1974), followed by coupling with 4-vinylpyridine (Friedman et al., 1970). Hydrolyses of 4-vinylpyridine-treated samples were performed in 4 M methanesulphonic acid or 3 M mercaptoethanesulphonic acid at 110 °C for 24 h as described above. N-terminal amino acid sequences were determined by automated Edman degradations using an Applied Biosystems Model 470A gas-phase sequencer (Thim et al., 1987).

Mass spectrometry analysis was performed using an API III LC/MS/MS system (Sciex, Thornhill, Ontario, Canada). The triple quadrupole instrument has a mass-to-charge (*m/z*) range of 2400 and is fitted with a pneumatically assisted electrospray (also referred to as ion-spray) interface (Bruins et al., 1987; Covey et al., 1988). Sample introduction was done by a syringe infusion pump (Sage Instruments, Cambridge, MA) through a fused capillary (75 μm i.d.) with a liquid flow rate set at 0.5−1 μL/min. The instrument *m/z* scale was calibrated with the singly-charged ammonium adduct ions of poly(propylene glycols) (PPGs) under unit resolution. The accuracy of mass measurements is generally better than 0.02%.

The hITF monomer and dimer forms were analyzed by peptide mapping followed by sequence and mass spectrometry analyses on isolated peptide fragments. One milligram of hITF monomer and dimer, respectively, was dissolved in 1 mL of 0.2 M Tris buffer, pH 6.0, containing 2 mM $CaCl_2$. Each peptide was digested by the addition of 60 μg of thermolysin (Sigma, P-1512) and incubation for 21 h at 70 °C. The reactions were stopped by the addition of 5 μL of 4 M HCl. The peptide mappings were carried out by the injection of the digest mixtures (250 μL) onto a Vydac 218TP54 column (4.6 × 250 mm) equilibrated at 30 °C at a flow rate of 1.5 mL/min with 0.1% (v/v) TFA. The concentration of acetonitrile in the eluting solvent was raised to 50% (v/v) over 100 min. The absorbance at 214 nm was recorded, and peptide fragments corresponding to the UV peaks (Figure 6) were collected and subjected to sequence and mass spectrometry analyses.

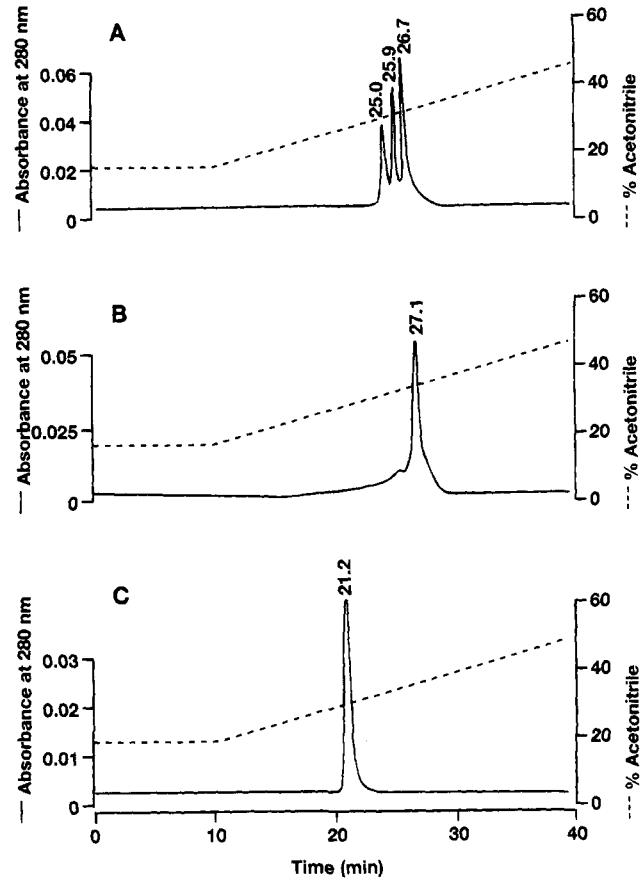

FIGURE 4: Reversed-phase HPLC on a Vydac 214TP54 C4 column of purified rat ITF (dimer) (A), human ITF (monomer) (B), and human ITF (dimer) (C). The dashed lines show the concentration of acetonitrile in the eluting solvent.

RESULTS

Expression and Purification. The expression levels for recombinant rat and human ITF in the present yeast system were 113 and 90 mg/L, respectively. This is the same level as we have previously found for recombinant hSP in the same system (Thim et al., 1993), but approximately 10 times higher than different insulin precursors in a similar system (Thim et al., 1986). The overall purification yields (24% for rITF and 50−65% for hITF) were also similar to the yields obtained for hSP. The loss during purification is mainly due to the complex medium in which the yeast cells are grown.

Characterization of rITF and hITF. Figure 4 shows the analytical HPLC chromatograms obtained on the purified rITF (Figure 4A) and hITF (Figure 4B,C). Recombinant rITF contains a mixture of three closely related peptides, and no attempts were made to separate these forms. When analyzed by electrospray mass spectrometry, three dominating molecular weights were found, corresponding to 13112.2, 13096.6, and 13078.8 (Figure 5A). The calculated molecular weight of rat ITF in a monomer form in which Cys-57 contains a free -SH group is 6558.3. The calculated molecular weight of rat ITF in a dimer form in which an S−S bridge is established between two Cys-57 residues is 13114.6. From the molecular weights found for the recombinant rat ITF, it is clear that all three peptides represent dimer forms of rITF. From other trefoil peptides in which the N-terminal amino acid residue is Gln, e.g., PSP (Thim et al., 1985; Tomasetto et al., 1990), it is known that this residue has a tendency to cyclize to form a pyrrolidone

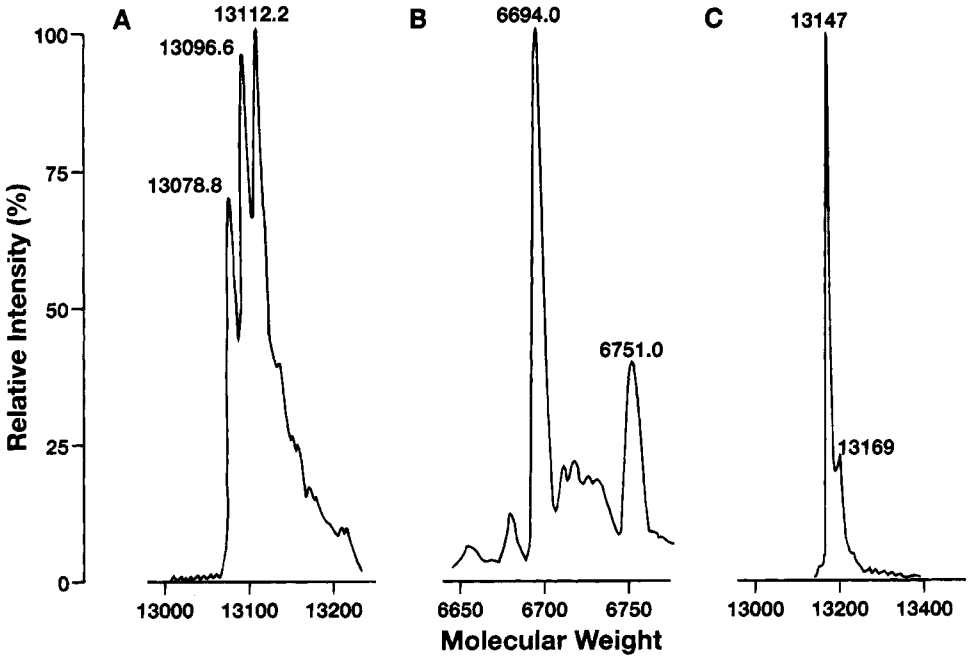

FIGURE 5: Reconstructed mass spectra of purified rat ITF (dimer) (A), human ITF (monomer) (B), and human ITF (dimer) (C).

carboxylic acid (pyrGlu). In the case of rat ITF having a predicted N-terminal sequence of Gln-Glu-Phe-Val-Gly-, it seems reasonable to assume that the N-terminal Gln could also cyclize to form pyrGlu. Such a derivatization would result in a decrease in the molecular weight of rat ITF (dimer) on 17 (one pyrGlu) or 34 (two pyrGlu) mass units, respectively. The observed molecular weights on 13096.6 and 13078.8 (Figure 5A) correspond to the dimer form of rat ITF in which one and two, respectively, N-terminal Gln residues have cyclized. The calculated molecular weight of these forms are 13097.6 and 13080.6, being a good agreement with the experimentally determined values. Thus from the HPLC (Figure 4A) and mass analysis (Figure 5A), it is assumed that the recombinant rat ITF consists of three different dimer forms: one containing two N-terminal Gln, one containing one N-terminal Gln and one N-terminal pyrGlu, and one containg 2 N-terminal pyrGlu. Table 1 shows the amino acid composition of rat ITF being in good agreement with the expected values.

Figure 4B,C shows the purity of hITF (monomer) and hITF (dimer), respectively, as analyzed by analytical HPLC. The dimer form (Figure 4C) looks relatively pure whereas the monomer form (Figure 4B) seems to be contaminated with material eluting in front of the peptide. However, upon rechromatography of material eluting in the main peak, a similar chromatogram was obtained (results not shown). This seems to indicate an atypical behavior of the hITF (monomer) on reverse-phase columns rather than impurities. We have previously observed a similar behavior of highly purified porcine PSP as well as highly purified recombinant hSP (Thim et al., 1993).

Mass spectrometry analysis of the hITF (monomer) shows a molecular weight of the dominating peak corresponding to 6694.0 (Figure 5B). The molecular weight, as calculated from the amino acid sequence (Figure 1), is 6574.4 assuming that Cys-57 exists on -SH form. The amino acid sequence analysis (Table 2) shows the expected N-terminal sequence of Glu-Glu-Tyr-Val-Gly-. The amino acid composition analysis (Table 1) shows the expected values except for the presence of 7.36 (8) cysteines. An additional cysteine linked to Cys-57 of hITF monomer would increase the molecular weight to 6694.7, which is very close to the value determined by mass spectrometry (6694.0). It is therefore assumed that in the hITF (monomer), Cys-57 is disulfide-linked to an additional cysteine. The minor molecular weight peak in the mass spectrum (Figure 5B) may represent another derivative of Cys-57 or may be an impurity in the preparation.

The calculated molecular weight of hITF (dimer), in which two monomers are linked by a disulfide bond between two Cys-57 residues, is 13146.8. This is in good agreement with the value determined by mass spectrometry (13147, Figure

Table 1: Amino Acid Composition of Rat ITF (Dimer), Human ITF (Monomer), and Human ITF (Dimer)[a]

amino acid	rat ITF (dimer)	human ITF (monomer)	human ITF (dimer)
Asx	11.92(12)	6.00(6)	12.01(12)
Thr[b]	8.00(8)	2.03(2)	4.01(4)
Ser[b]	9.86(10)	2.04(2)	4.01(4)
Glx	15.98(16)	6.92(7)	14.00(14)
Pro[c]	12.48(12)	5.97(6)	12.14(12)
Gly	6.02(6)	4.20(4)	8.09(8)
Ala	2.04(2)	3.91(4)	7.90(8)
Val	11.92(12)	4.92(5)	9.86(10)
Met	1.80(2)	0.00(0)	0.00(0)
Ile	1.98(2)	1.17(1)	2.13(2)
Leu	3.92(4)	2.03(2)	3.99(4)
Tyr[c]	2.02(2)	1.94(2)	3.92(4)
Phe	7.88(8)	2.93(3)	5.94(6)
Lys	2.14(2)	3.07(3)	6.09(6)
His	0.00(0)	0.99(1)	2.03(2)
Trp	2.16(2)	0.96(1)	1.79(2)
Arg	3.94(4)	2.96(3)	6.05(6)
PE-Cys[c]	12.70(14)	7.36(8)[d]	13.74(14)
total	(118)	(60)	(118)

[a] Values in parentheses are those deduced from cDNA: rat ITF (Suemori et al., 1991; Chinery et al., 1992) and human ITF (Hauser et al., 1993). [b] Determined by extrapolation to zero time of hydrolysis. [c] Determined by hydrolysis in 4 M methanesulfonic acid. [d] The hITF (monomer) contains eight cysteine residues, assuming that Cys-57 is disulfide linked to a cysteine.

Table 2: Automated Edman Degradation of Rat ITF (Dimer), Human ITF (Monomer), and Human ITF (Dimer)

	rat ITF (dimer)		human ITF (monomer)		human ITF (dimer)	
cycle no.	PTH-AA	yield (pmol)	PTH-AA	yield (pmol)	PTH-AA	yield (pmol)
1	Gln	989	Glu	1517	Glu	2227
2	Glu	731	Glu	1998	Glu	2361
3	Phe	967	Tyr	2205	Tyr	2528
4	Val	1072	Val	2409	Val	2431
5	Gly	701	Gly	1625	Gly	1803
6	Leu	838	Leu	2318	Leu	2253
7	Ser	497	Ser	852	Ser	904
8	Pro	965	Ala	1729	Ala	1712
9	Ser	406	Asn	1394	Asn	1518
10	Gln	820	Gln	1494	Gln	1499
11	(Cys)	n.d.	(Cys)	n.d.	(Cys)	n.d.
12	Met	581	Ala	1243	Ala	1377
13	Val	971	Val	1468	Val	1356
40	Pro	962	Pro	1223	Pro	1195
15	Ala	529	Ala	1248	Ala	1249
16	Asn	791	Lys	1270	Lys	1050
17	Val	952	Asp	937	Asp	991
18	Arg	331	Arg	891	Arg	964
19	Val	956	Val	1002	Val	1069
20	Asp	476	Asp	847	Asp	932
21	(Cys)	n.d.	(Cys)	n.d.	Cys	n.d.
22	Gly	381	Gly	709	Gly	703
23	Tyr	352	Tyr	894	Tyr	792
24	Pro	927	Pro	766	Pro	701
25	Thr	498	His	309	His	236
26	Val	812	Val	755	Val	670
27	Thr	510	Thr	505	Thr	576
28	Ser	225	Pro	458	Pro	473
29	Glu	398	Lys	444	Lys	321
30	Gln	499	Glu	219	Glu	304
31	(Cys)	n.d.	(Cys)	n.d.	(Cys)	n.d.
32	Asn	557	Asn	312	Asn	300
33	Asn	591	Asn	464	Asn	503
34	Arg	196	Arg	325	Arg	294
35	Gly	222	Gly	239	Gly	223
36	(Cys)	n.d.	(Cys)	n.d.	(Cys)	n.d.
37	(Cys)	n.d.	(Cys)	n.d.	(Cys)	n.d.
38	Phe	335	Phe	189	Phe	174
39	Asp	275	Asp	133	Asp	137
40	Ser	164	Ser	49	Ser	43
RY(av)		97.0%		93.2%		92.5%

5C). The other peak in the mass spectrum (13169) probably represents the Na^+ adduct of hITF (dimer). The sequence analysis (Table 2) as well as the amino acid composition analysis (Table 1) are also in good agreement with the expected values.

In order to confirm that dimer formation occurs by the establishment of a disulfide bond between Cys-57 of two monomers, comparative peptide mapping of the hITF (monomer) and hITF (dimer) was carried out (Figure 6). As can be seen from these results, the two peptide maps look very similar. However, the peptide map of the hITF (monomer) contains one peak (Figure 6, peak 1) that is absent in the peptide map of the hITF (dimer), and the peptide map of the hITF (dimer) contains two peaks (Figure 6, peaks 4 and 6) that are absent in the peptide map of the hITF (monomer). Mass spectrometry analyses of peptide fragments corresponding to these three peaks are given in Table 3. By combination of sequence analyses data and mass spectrometry data, the nature of these three peptides can be deduced (Table 3). Thus, in the structure of the hITF (monomer), Cys-57 is linked by a disulfide bond to a free

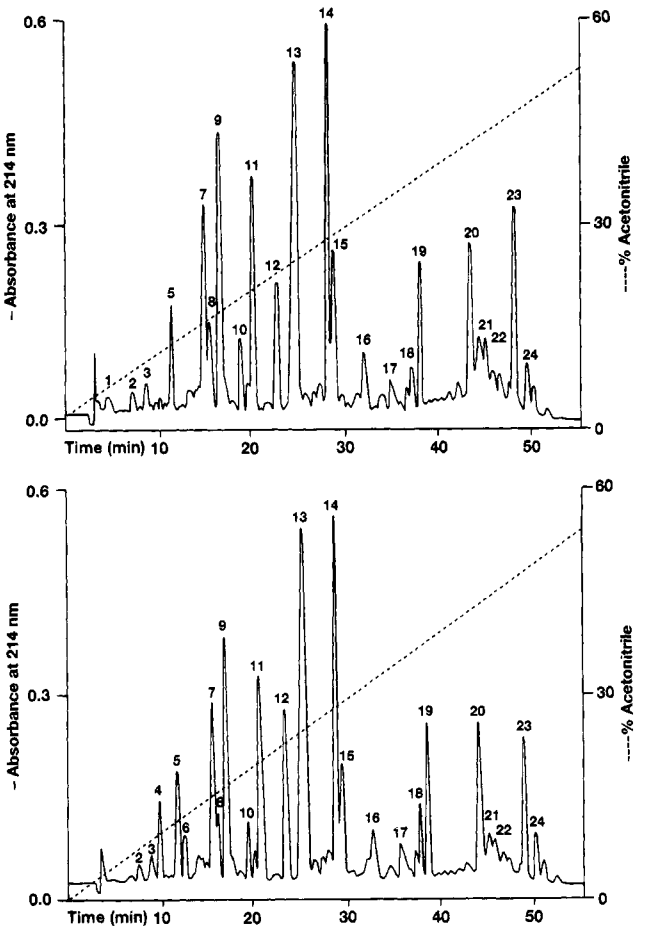

FIGURE 6: Peptide map of human ITF (monomer) (A) and human ITF (dimer) (B). Fractions corresponding to the individual absorbance peaks were collected as indicated. Notice the absence of peak 1 in the hITF (dimer) and the absence of peaks 4 and 6 in the hITF (monomer).

Table 3: Mass Spectrometry Analyses of Thermolysin Fragments of hITF (Monomer) and hITF (Dimer)

			MW theory		MW found	
peak no.[a]	origin	peptide	un-reduced	re-duced	un-reduced	re-duced
1	hITF (monomer)	Ala-Glu-Cys	440.5	321.3	440.0	no[b]
		Cys		121.2		no[b]
4	hITF (dimer)	Ala-Glu-Cys	569.6	321.3	569.8	322.0
		Glu-Cys		250.3		no[b]
6	hITF (dimer)	Ala-Glu-Cys	640.7	321.3	639.8	322.0
		Ala-Glu-Cys				

[a] Peak no. refers to the peptide maps (Figure 6). [b] no, not observed.

cysteine residue whereas in the hITF (dimer), two Cys-57 residues are disulfide linked to form the dimer (Figure 7).

DISCUSSION

In the present study, we describe the use of a yeast expression system for the production of milligram to gram amounts of recombinant rat ITF and human ITF. The expression level in this system is relatively high (~100 mg/L), and the system thus seems to be an alternative to the *Escherichia coli* system described by Chinery et al. (1993) in which the expression level was significantly lower (~100

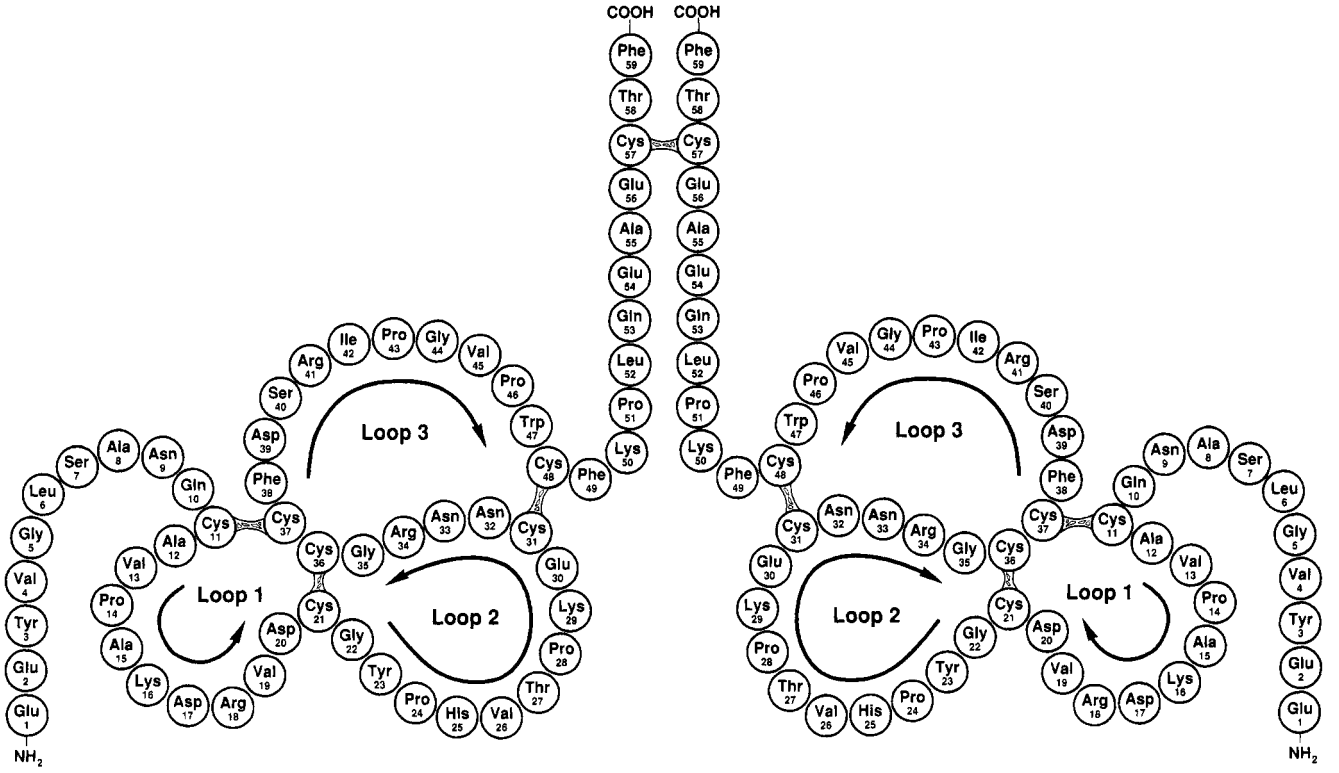

FIGURE 7: Structure of human ITF (dimer).

μg/L). Furthermore, the purification is relatively simple, resulting in an overall yield of 25−65% of the correctly processed peptide compared to the *E. coli* system in which the purification yield is lower and the peptide is obtained as an unfolded β-galactosidase fusion protein that has to be further processed and folded to obtain the native trefoil peptide (Chinery et al., 1993). The expression levels obtained for rat and human ITF in the present study are very similar to what we have previously reported for hSP (Thim et al., 1993). In the case of yeast expressed hSP, this peptide was found in the N-glycosylated form and in an unglycosylated form (Thim et al., 1993). Since neither rat nor human ITF contain the consensus sequence Asn-X-Thr or Asn-X-Ser, no potential N-glycosylation sites are present in these trefoil peptides. From the mass analysis of the purified rat and human ITF, it is also clear that no O-glycosylations occur in the yeast-expressed rat and human ITF.

Neither rat ITF nor human ITF have been purified and characterized from natural sources. Thus, the nature of any naturally occurring posttranslational modifications is unknown. In the case of human ITF, it is also unknown whether this peptide occurs in a monomer or in a dimer form. Preliminary gel filtration studies of rat ITF from intestinal goblet cells (Taupin & Giraud, 1994) and neoplastic colonic tissues (Taupin et al., 1994) seem to indicate that this peptide occurs in a 6.6-kDa form, thus existing as a monomer.

The yeast-expressed rat ITF was only found in a dimer form whereas human ITF was found both in monomer and in dimer form. The reason for this different behavior of two closely related peptides is not known. It could not be completely excluded that yeast-expressed rat ITF could also be present in a monomer form, which is removed during the purification of the dimer form. However, if this is the case, the monomer form would constitute only a few percent of the dimer form. For the yeast-expressed human ITF, the ratio between monomer and dimer was approximately 72−28%.

The yeast-expressed human ITF having the N-terminal sequence of Glu-Glu-Tyr-Val-Gly- was found not to form pyrrolidone carboxylic acid (both for the monomer and the dimer form). This is identical to the yeast-expressed hSP having an N-terminal sequence of Glu-Lys-Pro-Ser-Pro- in which the N-terminal Glu did not cyclize either (Thim et al., 1993). However, for the rat ITF having the N-terminal sequence of Gln-Glu-Phe-Val-Gly-, the N-terminal Gln was found partly to form pyrGlu, resulting in three forms of the rat ITF (dimer). This is comparable to the porcine PSP having the N-terminal sequence of Gln-Lys-Pro-Ala-Ala- (Tomasetto et al., 1990), which when isolated from porcine tissue was found in a form in which Gln was totally cyclized to pyrGlu (Thim et al., 1985). Both Glu and Gln can cyclize to form pyrGlu (Wold, 1981); however, the above results seem to indicate that the tendency is bigger for Gln than for Glu.

Preliminary attempts to digest native ITF with a series of proteolytic enzymes including pepsin, trypsin, chymotrypsin, and subtilisin have failed, confirming previous results (Jørgensen et al., 1983; Thim, 1989) that native trefoil peptides are extremely stable toward proteolytic digestion. In the present study, it has been possible to hydrolyze the native human ITF by the use of prolonged exposure of the molecule to thermolysin at 70 °C. Using peptide fragments purified from this digestion mixture, it has been proven that the hITF (dimer) is composed of two hITF monomers disulfide linked via two Cys-57 residues (Figure 7). Final proof for the position of the other disulfide bonds in ITF may come from three-dimensional structure analysis by X-ray (Gajhede et al., 1993; De et al., 1994) or NMR (Carr et al., 1994). Especially the monomer form of human ITF might be well suited for such studies. As both rat and human ITF

are difficult to purify in any significant amount from tissue extraction, the yeast expression system described in the present study may become a valuable tool in obtaining pure peptide for physiological, pathophysiological, and structural studies of trefoil peptides.

REFERENCES

Babyatsky, M. W., Thim, L., & Podolsky, D. K. (1994) *Gastroenterology 106*, A43 (abstract).

Bruins, A. P., Covey, T. R., & Henion, J. D. (1987) *Anal. Chem. 59*, 2642−2646.

Carr, M. D., Bauer, C. J., Gradwell, M. J., & Feeney, J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2206−2210.

Chinery, R., Poulsom, R., Rogers, L. A., Jeffery, R. E., Longcroft, J. M., Hanby, A. M., & Wright, N. A. (1992) *Biochem. J. 285*, 5−8.

Chinery, R., Poulsom, R., Elia, G., Hanby, A. M., & Wright, N. A. (1993) *Eur. J. Biochem. 212*, 557−563.

Covey, T. R., Bonner, R. F., Shushan, B. I., & Henion, J. D. (1988) *Rapid Commun. Mass Spectrom. 2*, 249−256.

De, A., Brown, D. G., Gorman, M. A., Carr, M., Sanderson, M. R., & Freemont, P. S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1084−1088.

Dignass, A., Lynch-Devaney, K., Kindon, H., Thim, L., & Podolsky, D. K. (1994) *J. Clin. Invest. 94*, 376−383.

Friedman, M., Krull, L. H., & Cavins, J. F. (1970) *J. Biol. Chem. 245*, 3868−3871.

Gajhede, M., Petersen, T. N., Henriksen, A., Petersen, J. F. W., Dauter, Z., Wilson, K. S., & Thim, L. (1993) *Structure 1*, 253−262.

Hanby, A. M., Poulsom, R., Elia, G., Singh, S., Longcroft, J. M., & Wright, N. A. (1993) *J. Pathol. 169*, 355−360.

Hauser, F., & Hoffmann, W. (1991) *J. Biol. Chem. 266*, 21206−21309.

Hauser, F., & Hoffmann, W. (1992) *J. Biol. Chem. 267*, 24620−24624.

Hauser, F., Roeben, C., & Hoffmann, W. (1992) *J. Biol. Chem. 267*, 14451−14455.

Hauser, F., Poulsom, R., Chinery, R., Rogers, L. A., Hanby, A. M., Wright, N. A., & Hoffmann, W. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6961−6965.

Hoffmann, W. (1988) *J. Biol. Chem. 263*, 7686−7690.

Hoffmann, W., & Hauser, F. (1993) *Trends Biochem. Sci. 18*, 239−243.

Jakowlew, S. B., Breathnach, R., Jeltsch, J.-M., Masiakowski, P., & Chambon, P. (1984) *Nucleic Acids Res. 12*, 2861−2878.

Jørgensen, K. H., Jørgensen, K. D., & Thim, L. (1983) U.S. Patent 4.307.317.

Lefebvre, O., Wolf, C., Kédinger, M., Chenard, M.-P., Tomasetto, C., Chambon, P., & Rio, M. C. (1993) *J. Cell Biol. 122*, 191−198.

Playford, R. J., Marchbank, T., Chinery, R., Evison, R., Pignattelli, M., Bolton, R., Thim, L., & Hanby, A. M. (1995) *Gastroenterology 108*, 108−116.

Podolsky, D. K., Lynch-Devaney, K., Stow, J. L., Oates, P., Murgue, B., DeBeaumont, M., Sands, B. E., & Makida, Y. R. (1993) *J. Biol. Chem. 268*, 6694−6702.

Poulsom, R., & Wright, N. A. (1993) *Am. J. Physiol. 265*, G205−G213.

Poulsom, R., Chinery, R., Sarraf, C., Lalani, E.-N., Stamp. G., Elia, G., & Wright, N. (1992) *Gastroenterology 27* (Suppl. 192), 17−28.

Prud'homme, J. F., Fridlansky, F., Le Cunff, M., Atger, M., Mercier-Bodart, C., Pichon, M.-F., & Milgrom, E. (1985) *DNA 4*, 11−21.

Rio, M. C., Bellocq, J. P., Daniel, J. Y., Tomasetto, C., Lathe, R., Chenard, M. P., Batzenschlager, A., & Chambon, P. (1988) *Science 241*, 705−708.

Rio, M.-C., Chenard, M.-P., Wolf, C., Marcellin, L., Tomasetto, C., Lathe, R., Belocq, J.-P., & Chambon, P. (1991) *Gastroenterology 100*, 375−379.

Rüegg, U. Th., & Rudinger, J. (1974) *Isr. J. Chem. 12*, 391−401.

Sherman, F., Fink, G. R., & Hicks, J. B. (1981) *Methods in yeast genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Suemori, S., Lynch-Devaney, K., & Podolsky, D. K. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11017−11021.

Taupin, D. R., & Giraud, A. S. (1994a) *Gastroenterology 106*, A635 (abstract).

Taupin, D. R., Ooi, K., Shulkes, A., & Giraud, A. S. (1994b) *Gastroenterology 106*, A637 (abstract).

Thim, L. (1989) *FEBS Lett. 250*, 85−90.

Thim, L., Jørgensen, K. H., & Jørgensen, K. D. (1982) *Regul. Peptides 3*, 221−230.

Thim, L., Thomsen, J., Christensen, M., & Jørgensen, K. H. (1985) *Biochim. Biophys. Acta 827*, 410−418.

Thim, L., Hansen, M. T., Norris, K., Hoegh, I., Boel, E., Forstrom, J., Ammerer, G., & Fiil, N. P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6766−6770.

Thim, L., Hansen, M. T., & Soerensen, A. R. (1987) *FEBS Lett. 212*, 307−312.

Thim, L. (1994) *Digestion 55*, 353−360.

Thim, L., Norris, K., Norris, F., Nielsen, P. F., Bjoern, S. E., Christensen, M., & Petersen, J. (1993) *FEBS Lett. 318*, 345−352.

Tomasetto, C., Rio, M., Gautier, C., Wolf, C., Hareuveni, M., Chambon, P., & Lathe, R. (1990) *EMBO J. 9*, 407−414.

Wold, F. (1981) *Annu. Rev. Biochem. 50*, 783−814.

Wright, N. A., Pike, C., & Elia, G. (1990) *Nature 343*, 82−85.

Wright, N. A., Poulsom, R., Stamp, G., Van Norden, S., Sarraf, C., Elia, G., Ahnen, D., Jeffery, R., Longcroft, J., Pike, C., Rio, M.-C., & Chambon, P. (1993) *Gastroenterology 104*, 12−20.

BI942373G