

CLONING AND SEQUENCING OF A cDNA ENCODING FOR A PROTEIN SPECIFICALLY EXPRESSED IN HUMAN TESTIS, OVARY AND PLACENTA

Rajesh K. Naz, Xiaolong Zhu and Changanamkandath Rajesh

Division of Research, Department of Obstetrics and Gynecology, Health education Building, Rm. 211, Medical College of Ohio, 3055 Arlington Avenue, Toledo, OH 43614-5806, USA

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Materials and methods
 - 3.1. Differential display technique to obtain testis-specific fragment
 - 3.2. Library screening, and isolation, sequencing and analysis of cDNA
 - 3.3. Rapid amplification of 5' and 3' cDNA ends (RACE) to obtain full-length cDNA
 - 3.4. Northern blot analysis
 - 3.5. Reverse transcription-polymerase chain reaction (RT-PCR)-southern blot procedure
4. Results and Discussion
5. Acknowledgement
6. References

1. ABSTRACT

Reproductive cell/tissue-specific antigens are attractive candidates for the development of a contraceptive vaccine. Using the differential display technology, a human cDNA fragment of 322-bp, designated as T17, was identified showing specific expression in the human testis. The T17 cDNA fragment was used as a probe to screen the human testis cDNA- λ gt10 library. After screening, one positive clone of ~1.1 kb having T17 nucleotide (nt) sequence was obtained. The 5' and 3' termini of this cDNA clone were extended by using 5' and 3' RACE procedures that yielded a full-length cDNA, designated as the TOP gene. The TOP cDNA is 1,480-bp long and has an ORF of 463 aa with the first ATG Met start codon at nucleotide (nt) 64 and the stop codon TGA at nt 1452. The translated protein has a calculated molecular mass of 49.3 kD with isoelectric point of 12.17. The deduced amino acid (aa) has one potential N-glycosylation site, and several phosphorylation and myristoylation sites. Hydrophathy plot generated from the deduced aa sequence showed it to be a surface protein. Comprehensive computer search in the database did not reveal any homology to any existing sequences both at the nt and aa levels. The TOP cDNA was found to be completely localized on the human chromosome 16 at the nt position 507833-506354, with the TOP ORF at the nt position 507770-506354. Northern blot analysis using three human Northern blots, indicated the specific expression of TOP gene in the human testis, ovary and placenta. RT-PCR-Southern blot analysis also confirmed tissue-specific expression of the TOP gene. The TOP cDNA may help us to gain insight into transcriptional control of the differentially expressed reproductive tissue-specific genes. It may also find clinical applications in the development of a contraceptive vaccine, and specific diagnosis and treatment of infertility in humans.

2. INTRODUCTION

The human population has exceeded over 6.2

billion, and is affecting growth and development in every sector universally. Although teenage pregnancy rates have declined in the USA, half the pregnancies are still unplanned (1), resulting in over a million elective abortions each year (2). Majority of these women were using some type of contraception, thus demanding better methods of birth control. Contraceptive vaccine (CV) has been proposed as an alternative method. It may provide a viable and acceptable birth control modality that will be inexpensive, long-lasting, inconspicuous, reversible and will require infrequent administration. Active immunization of various species of animals with extracts of sperm/testis or ovary results in infertility (3-5). However, the extracts of whole sperm/testis or ovary cannot be used for the development of a CV, because these cells/tissues have several antigens that could be shared with various somatic cells (6). The utility of an antigen in the development of a CV is contingent upon its specific expression in the reproductive cells/tissues and not in the somatic cells and tissues. Ours and various other laboratories are actively engaged in seeking such antigens, that are exclusively expressed in sperm/testis (7-10) or ovary (11,12), which can be used for immunocontraception. Complementary DNAs of a few of these antigens have been cloned and sequenced, and the effects of the encoded recombinant proteins have also been investigated on fertility (13-16).

The DeltaTM differential display technique (DDT) is a powerful approach to identify differentially expressed cDNAs in various cells and tissues (17-20). Our laboratory has used this technology to clone and sequence testis-specific (21) as well as prostate-specific genes (22, 23). In order to search for additional tissue-specific antigens, we are continuing to use this technology. In the present study, we report a cDNA that is specifically expressed in the human testis, ovary, and placenta. This cDNA was obtained originally by DDT and was found to be novel, without any homology with any known amino acid (aa) and nucleotide

Human testis, ovary and placenta specific cDNA

(nt) cDNA sequence in the database. The selected cDNA fragment was extended to a full-length cDNA by screening the λ -gt10 human testis library and then by using the 5' Rapid Amplification cDNA Ends (RACE) and 3' RACE procedures. The cDNA and its translated ORF were analyzed for any homology in the database, and the ORF was also examined for hydrophilicity and various potential glycosylation, phosphorylation and myristoylation sites. The tissue specificity was analyzed by the Northern blot analysis and confirmed by the Reverse Transcription-Polymerase Chain Reaction (RT-PCR)-Southern blot procedure. The long-term goal of this study is to search for reproductive cell/tissue-specific antigens that may have potential applications in the development of a CV, and also in the specific diagnosis and treatment of infertility in humans.

3. MATERIALS AND METHODS

3.1. Differential display technique to obtain testis-specific fragment

Total RNA of human testis, liver and kidney and Delta™ differential display kit were purchased from Clontech (Palo Alto, CA). The first-strand cDNA was synthesized by mixing 2 microgram of the total RNA of human testis, liver and kidney with 1 μ l of synthesis primer (1 μ M) in 5 μ l volume. The contents were incubated at 70°C for 3 min, then 2 μ l of 5x first strand buffer, 2 μ l of dNTP (5 mM), 1 μ l of Molony murine leukemia virus (MMLV) reverse transcriptase (200 units/ μ l), were added and the mixture was incubated at 42°C for 1 h. The first strand cDNA was polymerase chain reaction (PCR)-amplified using different P/T primer pairs provided with the kit. The reaction cocktail was prepared as follows: 2 μ l of 10X Ken Taq reaction buffer, 14.2 μ l of sterile water, 0.2 μ l of dNTP (5 mM), 0.2 μ l of [α -³²P] dATP, 0.4 μ l of advantage Ken Taq polymerase, in a total volume of 20 μ l. The PCR thermoprofile was: 94°C for 5 min, 40°C for 5 min, 68°C for 5 min, 1 cycle; 94°C for 2 min, 40°C for 5 min, 68°C for 5 min, 2 cycles; and 94°C for 1 min, 60°C for 1 min, 68°C for 2 min, 25 cycles.

PCR products were resolved by 5% polyacrylamide/8 M urea gel electrophoresis the gel was dried and exposed to an X-ray film. After aligning the positive and negative controls on the autoradiograph, the testis-specific bands were identified, cut out, dissolved in sterile water, heated at 100°C for 5 min, centrifuged, and the supernatants were collected for further analysis. Each differentially-displayed band was re-amplified using the corresponding primers in the original PCR procedure. The thermoprofile was: 94°C for 1 min, 60°C for 1 min, 68°C for 2 min, 35 cycles. The PCR products were separated by 1% agarose gel electrophoresis, the desired bands (same size as the original band) were eluted, purified, cloned into pBluescript II SK+ vector at *Sma* I site, and then sequenced with T3 and T7 promoter primers using the dideoxy chain-termination method (24) employing sequenase version 2.0 DNA sequencing kit (Amersham Life Science Inc., Buckinghamshire, UK). Several P/T primer combinations were tried. The testis-specific T17 cDNA fragment

described was detected by using the P2/T7 primer combination. The nucleotide sequence of T17 was analyzed in the database for homology with any known cDNA sequence.

3.2. Library screening, and isolation, sequencing and analysis of cDNA

The human testis cDNA library in λ -gt10 (Clontech) was plated at a density of 10×10^5 plaque-forming units (pfu) per 100-mm Petri dish with *E. coli* C600 as host bacterium. The plates were incubated at 37°C until the plaques reached a diameter not exceeding 1.5 mm. The nitrocellulose filters were placed on the plates for 60 sec, and then taken off, denatured, neutralized, and followed by UV crosslinking. The probe used for screening the library was the T17 cDNA fragment labeled with ³²P-dCTP by using the random hexamer method. The plaques of the putative positives clones were amplified, and their DNAs were extracted, purified and cut by *Eco* RI. The inserts were eluted from the agarose gel, purified, subcloned into pBluescript II SK+ at *Eco* RI site, and sequenced by using T3 and T7 promoter primers and several internal primers based on the obtained sequences. The cDNA sequence was aligned and translated using the software Macvector DNA Analysis (Oxford Molecular group, Campbell, CA). Hydrophilicity plot of deduced amino acid sequence of the Open Reading Frame (ORF) was analyzed by using the Kyte and Doolittle (1982) (25) and the Engelmann *et al.* (1986) (26) methods. The transmembrane characteristic of the amino acid sequence was analyzed by web-based transmembrane prediction server program (<http://www.biokem.su.se/~server/DAS/tmdas.cgi>). The nt and aa sequence homology search was performed by BLASTn and BLASTx algorithm (27) and GCG program on the National Center for Biotechnology information server (<http://www.ncbi.nlm.nih.gov>).

3.3. Rapid amplification of 5' and 3' cDNA ends (RACE) to obtain full-length cDNA

Since the cDNA obtained from the λ -gt10 library was incomplete, the RACE procedure was performed to obtain the 5' and 3' termini (Clontech, CA). Briefly, the poly (A)⁺ RNA of human testis was used to synthesize the first strand cDNA using MMLV reverse transcriptase, followed by the second strand synthesis using DNA polymerase as described elsewhere [21]. Marathon cDNA adapter (Clontech) was ligated to the double-stranded DNA to generate the adapter-ligated cDNA library of human testis. 5'-RACE procedure was conducted by using the PCR procedure. The primers were: AP1 provided by the Clontech, and the gene-specific primer 1 (GSP1) at nt position 250-226 (5'-TGCACGGGTCGGGCGGTCTGCCTG-3') of the insert sequence. The template was the adapter-ligated cDNA library of human testis. The thermoprofile was: 94°C for 1 min, than 94°C for 30 sec, 68°C for 2 min, 30 cycles. The gene-specific band was identified using the Southern blot procedure, employing the gene-specific primer 2 (GSP2) at nt position 220-202 (5'-CGTGGGCGCGAAGGTTTCG-3') of the insert sequence. Briefly, the PCR product(s) was separated by 1% agarose gel electrophoresis, transferred

Human testis, ovary and placenta specific cDNA

onto nitrocellulose membrane, and hybridized (60°C, 60 min) with the P³²-labeled GSP2. The membrane was washed (10 min x 3), and exposed to the X-ray film. The gene-specific band was eluted from the agarose gel, purified by phenol/chloroform, and precipitated by ethanol. The DNA fragments were blunted, and cloned into pBluescript II SK+ vector at the *Sma* I site. The colony containing the insert was picked up and sequenced. The sequence indicated a ~200 bp sequence upstream of 5' end of the T17 cDNA, having one termination codon before the ATG start codon. 3'-RACE procedure was similar to the 5'-RACE procedure except differences in the gene-specific primers used in the PCR procedures. For the 3'-RACE procedure, the gene-specific primer 3 (GSP3) localized at nt position 1187-1210 (5'-CGGCGGGATCCTCCCCGACTCGG-3') of the cDNA and the AP1 primer were used for PCR reaction. The gene-specific primer 4 (GSP4) localized at nt position 1213-1230 (5'-GGGGAGGCGCGGGCCACA-3') of the cDNA was used for examining the authenticity of the amplified cDNA by using the Southern blot procedure. The 3'-RACE showed a ~300-bp sequence downstream of the 3' end of the T17 cDNA, having one stop codon. The full-length sequence will henceforth be referred to as testis, ovary and placenta (TOP) cDNA.

3.4. Northern blot analysis

The coding region of the full-length cDNA (ORF) was amplified by PCR with the TOP specific primers. The 5' primer, localized at the nt position 64-83 of the TOP cDNA, was 5'-ATGGGCGCGGCAGGGCGGGC-3' and the 3' primer, localized at the nt position 1452-1433, was 5'-TCACTGCAACCTCCAATTCC-3'. The template used was the human genomic DNA (Clontech, CA) since the TOP sequence was found to be localized completely on the human chromosome 16. The thermoprofile was: 94°C for 1 min, 50°C for 2 min, 72°C for 2 and half min, 35 cycles. The PCR products were separated by 1% agarose gel electrophoresis and the desired band of ~1.4 kb, was eluted from the gel, purified, and used as a probe. Two human Multiple Tissue Northern (MTN)TM blots were obtained from Clontech. The blots were incubated (60°C, 30 min) with QuickHyb solution (Stratagene, CA), then hybridized (60°C, 2 h) with α^{32} PdCTP-labeled ~1.4 kb probe, washed, and exposed to X-ray film for 24 h to five days. After the blots were stripped off the TOP cDNA probe, they were re-hybridized with α^{32} PdCTP-labeled β -actin probe, washed and exposed to X-ray film.

A recently available Northern blot membrane named High Throughput Northern BlotTM, containing mRNAs of 30 human adult normal tissues was also tested (BioChain Institute, Hayward, CA). This mRNA blot contains 640 spots in 32 lanes and each lane is comprised of 20 spots representing mRNA of a single tissue. The 32 lanes represent 30 different human adult normal tissues from which mRNAs are isolated and the fractions containing different sizes of mRNA species are recovered. The first and lane 21 of the blot are duplicates of mRNAs from placenta, and lane 2 is empty. The recovered mRNA fractions are arrayed on positively charged nylon membranes and UV cross-linked. The 20 arrayed fractions

for each tissue has a range of different sizes of the transcripts. The blot was incubated (65°C, 30 min) with QuickHyb solution (Stratagene, CA), then hybridized (65°C, 2 hr) with ³²P-labeled PCR-amplified TOP probe, washed, and exposed to X-ray film for 48 h to five days. The blot was the stripped off the TOP cDNA, washed and re-hybridized with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) DNA probe at 65°C overnight, washed, and the reactive spots were detected by exposure to X-ray film. The spots detected on the blot were compared with the table provided by the manufacturer to calculate the size of the transcript.

3.5. Reverse transcription-polymerase chain reaction (RT-PCR)-southern blot procedure

To further examine the tissue-specificity of the TOP, the sensitive RT-PCR –Southern blot procedure was performed. The total RNAs from 12 types of human tissues, namely, testis, small intestine, heart, lung, mammary gland, trachea, skeletal muscle, brain, kidney, liver, uterus, and spleen, were purchased from Clontech. To prepare the first strand cDNA, 2 μ g of the total RNA from each tissue was mixed with 0.5 μ g (0.5 mg/ml) of oligo(dT)₁₅ primer and 4 μ l of 5x buffer (250 mM Tris-HCl, pH 8.3, 3.75 mM KCl, 15 mM MgCl₂), and the mixture was heated at 70°C for 2 min and cooled on ice for 2 min. To this reaction mixture, 0.5 μ l (38 U/ μ l) of RNasin inhibitor, 2 μ l of 100 mM DTT, 1 μ l of 10 mM dNTPs, and 2 μ l (400 U) of MMLV reverse transcriptase were added. The reaction mixture was incubated at 37°C for 60 min, and then stored at -20°C. Two microliters of the first strand cDNA were amplified by PCR for 35 cycles (94°C for 1 min, 54°C for 2 min, 72°C for 2 min) by using the TOP specific primers. The 5' primer was 5'-CGGCCAGCCCCCGGAACCCT-3' localized at nt position 414-433 and the 3' primer was 5'-GCCGGCGTCCTAGCGGTTGG-3' localized at nt position 1161-1180. This pair of primers is expected to amplify a fragment of 766 bp. The β -actin specific primers of 29-mer were based on the conserved regions in the rat and human β -actin cDNA sequences; sense primer: 5'-GGGGATCCGTACATGGCTGGGGTGTGAA-3' and the antisense primer: 5'-GGGGATCCGGCGACGAGGCCAGAGCAAG-3' (28). The thermoprofile to PCR amplify the first-strand cDNAs with the β -actin primers was: 94°C for 1 min, 50°C for 1 min, 72°C for 1 min, 35 cycles. This primer set is expected to amplify a 256-bp fragment. PCR-amplified products were analyzed by the agarose gel electrophoresis, and the cDNAs were transferred to nitrocellulose membranes for the Southern blot procedure to examine the specificity of the amplified fragments. The membranes were prehybridized (60°C, 15 min) with QuickHyb solution, and then incubated (60°C, 2 hr) with ³²P-labeled TOP probe or β -actin probe. The membranes were washed and exposed to X-ray film for 1-2 days.

4. RESULTS AND DISCUSSION

Using DeltaTM differential display technology a human testis-specific 322-bp fragment, designated as T17,

Human testis, ovary and placenta specific cDNA

was identified employing P2/T7 primer pair. The T17 cDNA fragment was subcloned into pBluescript II SK+ vector and sequenced. The sequence analysis revealed that the cDNA was novel without any homology to any known cDNA sequence in the database. The 322 bp cDNA fragment was used as a probe to screen the human testis cDNA- λ gt10 library. Upon screening $\sim 5 \times 10^5$ pfu from the library, three putative positive clones were identified. After subcloning them into pBluescript II SK+ and sequencing, one putative clone was retrieved which had an insert of ~ 1.1 kb with T17 sequence. The nucleotide sequence was established by total sequencing of both strands by using T3 and T7 promoter primers and several internal primers based on the obtained sequences. The authenticity of the positive clone was confirmed by the presence of 322-bp T17 sequence. This clone has an incomplete sequence and the 5' and 3' termini were completed by using 5' and 3' RACE procedures. The alignment of the sequences yielded a 1,480-bp cDNA, designated as TOP (Figure 1). Computer translation (GCG program, Madison, WI) of the TOP cDNA generated an open reading frame (ORF) of 463 amino acid (aa) (Figure 1). The deduced aa has calculated molecular mass of 49.3 kD with isoelectric point of 12.17. It has one potential N-glycosylation site at aa position 444-447; one potential cAMP- and cGMP-dependent protein kinase at aa position 406-409 (KRES), eight potential casein kinase II phosphorylation sites, and 15 potential protein kinase C phosphorylation sites (Figure 1). There are 12 potential myristoylation sites on the deduced aa sequence. The presence of several phosphorylation sites are important since the phosphorylation at tyrosine, serine and threonine residues has been shown to have an important role in the development of fertilization capacity of the human sperm (29).

The presence of several myristoylation sites suggest binding capacity of the protein to the plasma membrane (30). Hydropathy plots computed by the Kyte and Doolittle method (25) and Engelman *et al.* method (26), and the surface probability analysis also suggest that the TOP protein could be a surface protein (Figure 2, Panels A-C). There are small hydrophobic pockets interspersed in the hydrophilic core at the N-terminus, center and the C-terminus. The N-terminus has a short hydrophobic region flanking a hydrophilic stretch, which characterizes a potential signal peptide. According to the (-3,-1)-rule (31), the cleavage site is between G at aa position 23 and W at aa position 25, the length of the leader sequence is 24 aa.

Comprehensive computer search by GCG and BLAST showed that the TOP is a novel gene both at the cDNA and protein levels. The TOP cDNA was found to be completely localized on the human chromosome 16 at nt position 507833-506354, and the ORF at the nt position 507770-506382. The TOP cDNA also has a higher similarity with the human chromosome Y at nt position 1019498-1018672 and with a human testis clone (gi:14001276) in the Expressed Sequence Tag (EST) database.

Tissue-specific expression of the TOP gene at the mRNA level was examined by the Northern blot and RT-

PCR-Southern blot procedures. In the Northern blot procedure, one specific band of ~ 1.7 kb was observed in the testis lane and one specific band of ~ 1.55 kb in the placenta lane. No band was observed in any other 13 human tissues in the two MTN blots tested (Figure 3, Panels A and B). The integrity and equal loading amount of RNA were confirmed by the signal (~ 2.0 kb) of equal intensity obtained in all tissue lanes after hybridization with the β -actin probe (10). Tissue-specific expression was also observed by using the High Throughput Northern blot containing 30 human adult normal tissues. It is a novel research blot recently developed by BioChain Institute (CA). It contains purified mRNAs rather than whole RNAs, and is comprised of mRNAs from more tissues than those in other commercially available blots. For preparation of the blot, total RNAs were isolated from 30 different normal human tissues, mRNAs isolated, separated by gel electrophoresis and the fractions containing mRNA species of different sizes were recovered for each tissue. The fractions from each tissue were arrayed on a single lane as 20 spots at equal distance, with each spot representing a range of different sizes of transcripts. When probed with glyceraldehyde-3-phosphate dehydrogenase (GAPDH), each tissue lane showed spots of 1.4 kb of equal intensity (Figure 4, Panel A). When probed with TOP cDNA, the spots were visible only in the ovary, placenta and testis lanes and not in the other 27 tissues (Figure 4, panel B). There were two spots of ~ 1.1 kb and ~ 1.9 kb, respectively, in the ovary lane and two spots of ~ 0.8 kb and ~ 1.1 kb, respectively, in the placenta lane, and only one spot of ~ 1.7 kb in the testis lane. The duplicate placenta lanes (lane no. 1 and 21) showed similar reaction patterns and the control lane (no.2) did not show any binding with the probe. Because of the expression in these three reproductive tissues (testis/ovary/placenta), this cDNA has been designated as the TOP gene. It has different transcript sizes in the three tissues with ~ 1.7 kb transcript expressed only in the testis.

The tissue-specific expression of the TOP cDNA (ORF) was further confirmed by using the sensitive RT-PCR-Southern blot procedure. The TOP specific primers amplified the expected band of 766-bp only in the human testis cDNA and not in the cDNAs from other 11 types of human tissues (Figure 5, Panel A). This band was specifically recognized by the TOP cDNA probe in the Southern blot procedure (Figure 5, Panel B). When the reverse-transcribed cDNA was PCR-amplified by using beta-actin specific primers, the expected 256-bp band was obtained in all tissues tested (Figure 5, Panel C), that was specifically recognized by the β -actin cDNA probe in the Southern blot procedure (Figure 5, Panel D). The findings with the beta-actin probe indicate the integrity and equal loading of RNA from all tissues. The result of the RT-PCR-Southern blot procedure confirmed the findings of the Northern blot procedure.

The reproductive cell/tissue specific TOP cDNA may help to gain insights into the transcriptional control of reproductive cell/tissue-specific genes during spermatogenesis, oogenesis and trophoblast/placental growth and differentiation. The TOP protein may also find clinical applications in the development of a contraceptive

Human testis, ovary and placenta specific cDNA

```

CCA AGT AGG TAG AAA GCG GCC TCC AGG CCT CGC GAG GAT GAG CCC GGC GTC CCA 54
      TER
GTC GCG AGG ATG GGC GCG GCA GGG CGG GCG ACC CGG CTT GTG GGA GGG AGG GAG 108
      M G A A G R A T R L V G G R E 15
CAG CGC GGG GCG GTG GGA GGG GGG TGG TGG GGC GGC GAA CCG GAC ATC CCA TAC 162
      Q R G A V G G G W W G G E P D I P Y 33
ACC CAC ACT ACA GAA CAC CCC CCG ATG GGC TCA CCA CTC CCG AAC CTT CGC GCC 216
      T H T T E H P P M G S P L P N L R A 51
CAC GTG CGA ACA GGC AGA CCG CCC GAC CCG TGC ACG GCA GCT GCG AGG GAC CGG 270
      H V R T G R P P D P C T A A A R D R 69
CGG CCA CTC GGG CGT GGC GGG CGC GGG GCA GCC CAG ACG TTT GGG CAG CGA ACC 324
      R P L G R G G R G A A Q T F G Q R T 87
AGA GGC GGA CCG CGG TGC CCG GGA TCT CAC CGC CAG CGG CCT CCA AGC ATG AAG 378
      R G C G P R C P G S H R Q R P P S M K 105
GCG GCC CGC CGG CTC CCC CAC CGC CTC CGC CAT CGC GGC CAG CCC CCG GAA CCC 432
      A A P R L P H R L R H R G Q P P E P 123
TCT TCC CTG CGC GCG CTG CAC GCC GAC CCC ATA CCC TCC GGG CTC TCA CCA GGC 486
      S S L R A L H A D P I P S G L S P G 141
CCA TGC GGA GCG CCG CCG ACC TGG TCC GGA AGG CGC GCG CCC GGG GAC AGG GAC 540
      P C G A P P T W S G R R A P G D R D 159
AAC CGG CCA ACC AGT GGC CCG TGG CAG CGC CAC ACA AGG CGG AGC CGG GTT TGG 594
      N R P T S G P W Q R H T R R S R V W 177
TCC CAG AGG GGG CCA CCA CAG CCT AAG CTG GTG AGC CGC TCG GGG AGA GAG GAT 648
      S Q R G P P Q P K L V S R S G R E D 195
-----
ACG CGG GCG GGG AGG GGG GCA CAG ACA GGC AAG GCC AGG GAC CGC AAG GGC AAG 702
      T R A G R G A Q T G K A R D R K G K 213
GGC ACC CCG GAG CCC GCA GAG GGG CGG CTC GGG AAG AAA CCT CAG GCA AAG CCG 756
      G T R E P A E G R L G K K P Q A K P 231
GGC CAC CAG GAA AAC ACG GCC ACG GGA TCC CAC CGC CAC AGA TAC GAG GGA GGT 810
      G H Q E N T A T G S H R H R Y E G G 249
CCC GCG GCG CCC CGC CTA GGA AGC CGG ACG GCC CCT GGC ACC CAC CGA GAC CCG 864
      P A A P R L G S R T A P G T H R D P 267
-----
CCT CAC GAG CCT GGG TCC CGC CAT CGG GAC CCC GAA GCG ACC TCA GCC ACA AAC 918
      P H E P G S R H R D P E A T S A T N 285
CCA ACG CCA GGG CCA CGT TGC TGG TTT CTT GTC CAT CCT CCG ACA CTG TCA AGC 972
      P T P G P R C W F L V H P P T L S S 303
TCC GGG AGA CCG GCG CGC CCC CCA CTT GGG ATG CTT CCC ATG GCC AGG CAT CCC 1026
      S G R P A R P P L G M L P M A R H P 321
AAC CCC GTG CCA CGC AAA CGC GGT TGT CGG CAC CGG TCG CTG CTC CTC AGG GGA 1080
      N P V P R K R G C R H R S L L L R G 339
GCG GGT GGA GAG CCG GCT CGC AGC GGA GCG GGT CAC GCG CCG GAC GGA ACG CCT 1134
      A G G E P A R S G A G H A P D G T P 357
GGC ACA GCC ACC GCT CGC GCA GCC TCC CAA CCG CTA GGA CGC CGG CCC GTC CCG 1188
      G T A T A R A A S Q P L G R R P G P 375
GCG GGA TCC TCC CCC GAC TCG GAA GGG GGA GGC GCG GGC CAC ACA GTA GGT GAC 1242
      A G S S P D S E G G G A G H T V G D 393
-----
GAG CCG CCC TCG GTC CCC ACC GCG GAG GAC TCT CTC AAG AGA GAG TCG GTA AGA 1296
      E P P S V P T A E D S L K R E S V R 411
-----
GAC TCA TCA TTC TCT CCC GAA AGC AGT GAG GTG GAT GGT GGC TTG TGT TCC GAG 1350
      D S S F S P E S S E V D G G L C S E 429
-----
CTC CTG TGG TTT CAG TTG GCC GCG CGT AGA AGA GAG ATT TCC AAT GTT TCC AGA 1404
      L L W F Q L A A R R R E I S N Y S R 447
GAG GCG CGA GCC ACA GTC ATT GAA CCC AGG AAG TGG AGG TTG CAG TGA 1452
      E A R A T V I E P R K W R L Q *

```

Figure 1. Nucleotide and deduced amino acid (shown in bold) sequences of the full-length TOP cDNA. There is one termination codon (TER) before the start codon ATG at the 5' end. One potential N-linked glycosylation site is shown in box. The deduced amino acid sequence also indicates one cAMP-cGMP kinase site (*****), eight casein kinase II phosphorylation sites (-----) and fifteen protein kinase C phosphorylation sites (_____).

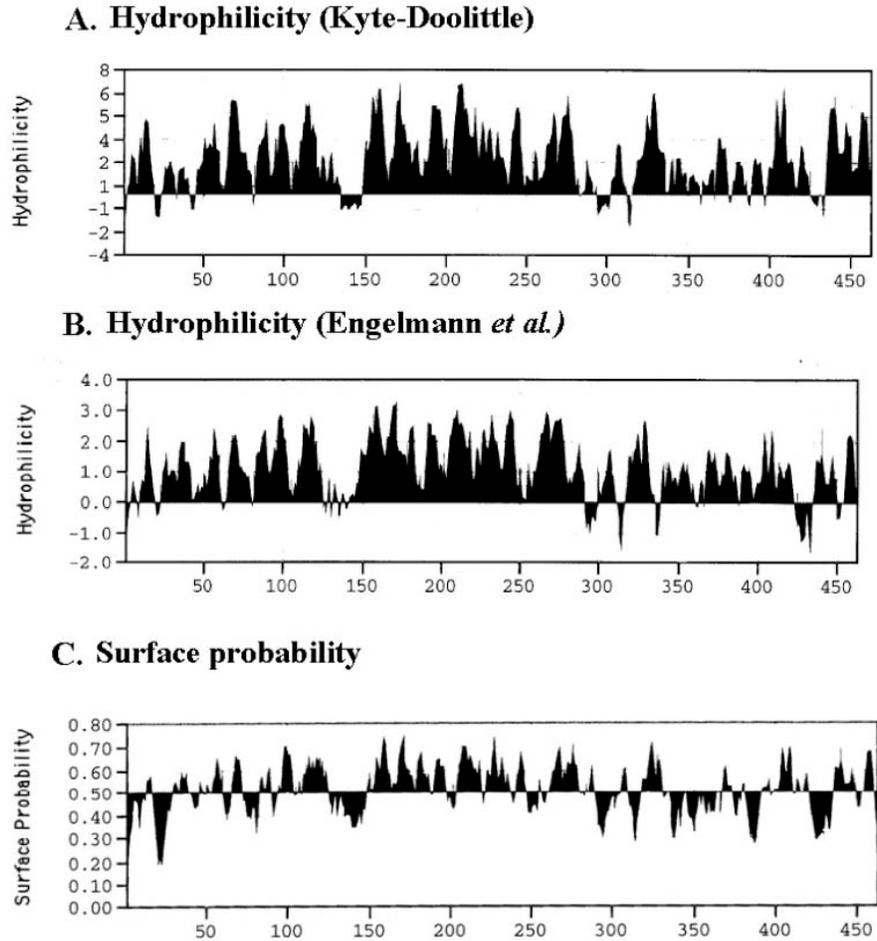


Figure 2. Hydropathy plots of the deduced aa sequence of TOP protein computed by the Kyte and Doolittle method (A) and Engelman *et al.* method (B). The surface probability analysis is shown in panel C.

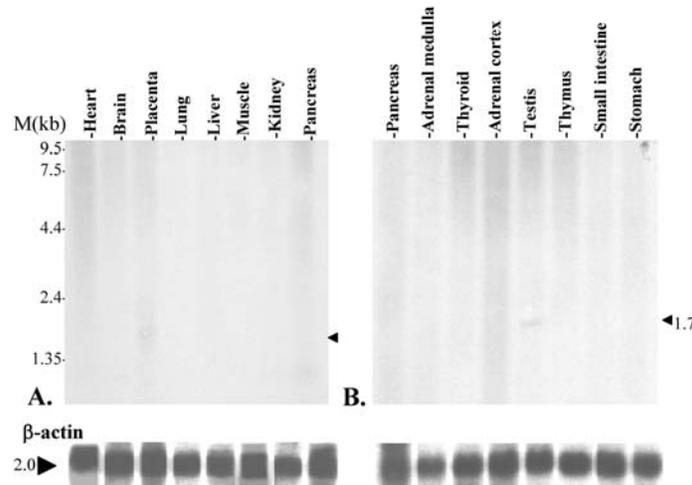


Figure 3. Tissue-specific expression of TOP gene in the Northern blot procedure. Two human multiple tissue blots (Panels A and B) containing RNAs from 15 human tissues were hybridized with the [α - 32 P] dCTP- labeled TOP gene (ORF), washed and exposed to X-ray film for 24 h. It showed a band of ~1.7 kb in the testis lane (arrowhead), a band of ~1.55 kb in the placenta lane (arrowhead), and not in any other 13 human tissues. The blots were stripped-off the TOP cDNA and re-hybridized with [α - 32 P] dCTP- labeled β -actin probe, that showed a band of ~2.0 kb of equal intensity in all lanes. The β -actin band shown in this Figure has been reproduced from an earlier work that utilized these blots (22).

Human testis, ovary and placenta specific cDNA

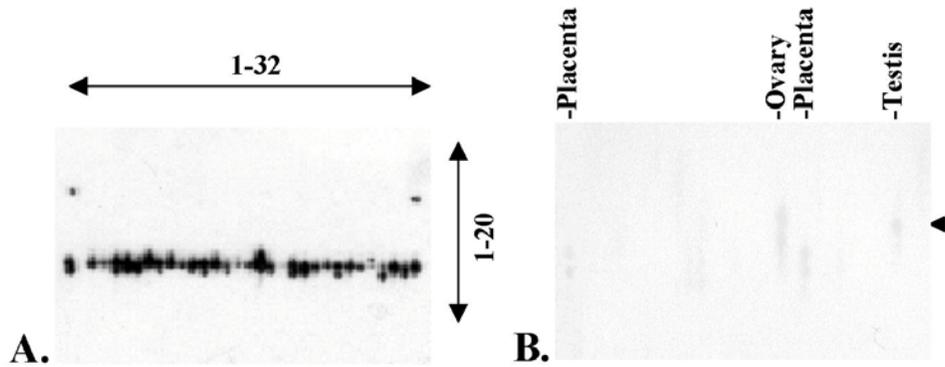


Figure 4. Tissue-specific expression of TOP gene on the High Throughput Northern Blot (BioChain Institute, Hayward, CA). The blot contains 30 human tissue mRNAs (1-32 lanes correspond to placenta (positive), blank, bladder, brain, frontal lobe brain, hippocampus brain, occipital lobe brain, parietal lobe, temporal lobe, breast, cerebellum, colon, esophagus, heart, kidney, liver, lung, muscle, ovary, pancreas, placenta, rectum, skin, small intestine/ duodenum, small intestine/jejunum, small intestine/jejunum, spleen, stomach, testis, tonsil, uterus/cervix, and uterus/corpus, respectively) arranged in 32 lanes with each lane having 20 spots of increasing size of transcripts from below to the top (1-20). (**Panel A**) when the blot was hybridized with the control Glyceraldehyde-3-phosphodehydrogenase (GAPDH) probe, each tissue lane showed a spot of ~1.4 kb of equal intensity. (**Panel B**) When probed with the TOP cDNA, the spots were visible only in the ovary, placenta and testis lanes (arrowhead). No spot was visible in any other tissue lane.

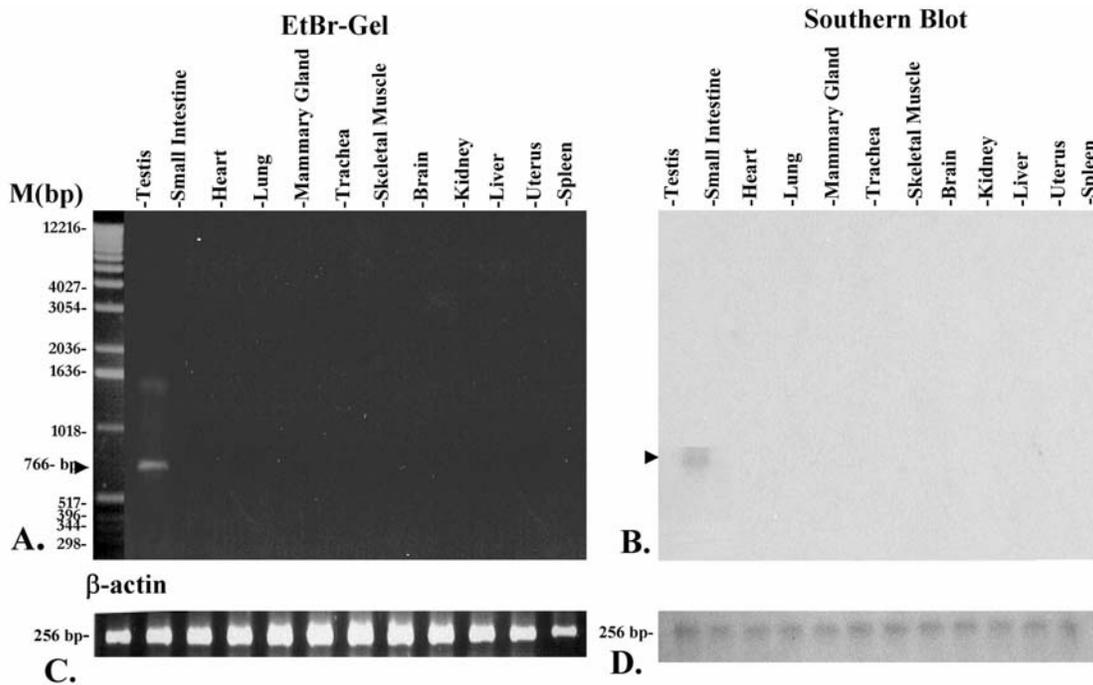


Figure 5. Tissue-specific expression of TOP gene in the RT-PCR-Southern blot procedure. (**Panel A**) Using the TOP-specific primers, the expected 766-bp fragment (arrowhead) was amplified only from the testis RNA and not from the RNAs of other 11 types of tissues tested in the ethidium bromide (EtBr)-stained agarose gel. (**Panel B**) The 766-bp amplified fragment was specifically recognized by the TOP cDNA probe in the Southern blot procedure. (**Panel C**) The β -actin primers amplified a 256-bp fragment from all the tissue RNAs that was specifically recognized by the beta-actin cDNA in the Southern blot analysis as shown in **Panel D**.

vaccine for men and women, and in the specific diagnosis and treatment of male and female infertility. Presently, we are expressing the TOP cDNA in appropriate vectors to obtain the recombinant protein that will be used to raise specific antibodies for examining its immunocontraceptive effects.

5. ACKNOWLEDGEMENT

This work was supported in part by NIH grant HD 24425 to RKN.

6. REFERENCES

1. Nelson, A: Contraceptive update Y2K: need for contraception and new contraceptive options. *Clin Cornerstone* 3, 48-62 (2000)
2. Grow, D. R. and S. Ahmed: New contraceptive methods. *Obstet Gynecol Clin North Am* vii-viii 27, 901-916 (2000)
3. Edwards, R. G: Immunological control of fertility in female mice. *Nature* 203, 50-53 (1964)
4. Menge, A. C. and R. K. Naz: Immunologic reactions involving sperm cells and preimplantation embryos. *Am J Reprod Immunol Microbiol* 18, 17-20 (1988)
5. Dunbar, B. S: Ovarian antigens and infertility. *Am J Reprod Immunol* 21, 28-31 (1997)
6. Naz, R. K. and A. Menge: Development of antisperm contraceptive vaccine for humans: why and how? *Hum Reprod* 5, 511-518 (1990)
7. Primakoff, P., W. Lathrop, L. Woolman, A. Cowan and D. Myles: Fully effective contraception in male and female guinea pigs immunized with the sperm protein PH-20. *Nature* 335, 543-546 (1988)
8. Lea, I. A., M. J. van Lierop, E. E. Widgren, A. Grootenhuis, Y. Wen, M. van Duin and M. G. O'Rand: A chimeric sperm peptide induces antibodies and strain-specific reversible infertility in mice. *Biol Reprod* 59, 527-536 (1998)
9. Naz, R. K: Vaccine for contraception targeting sperm. *Immunol Rev* 33, 1-32 (1999)
10. Naz, R. K. and X. Zhu: Molecular cloning and sequencing of cDNA encoding for human FA-1 antigen. *Mol Reprod Dev* 63, 256-268 (2002)
11. Hennebold, J. D., M. Tanaka, J. Saito, B. R. Hanson, and E. Y. Adashi: Ovary-selective genes I: the generation and characterization of an ovary-selective complementary deoxyribonucleic acid library. *Endocrinology* 141, 2725-2734 (2000)
12. Bagavant, H., F. M. Fusi, J. Baisch, B. Kurth, C. S. David and K. S. Tung: Immunogenicity and contraceptive potential of a human zona pellucida 3 peptide vaccine. *Biol Reprod* 56, 764-770 (1997)
13. Millar, S. E. and J. Dean: Targeting the zona pellucida for immunocontraception. In: *Immunology of Reproduction* Eds: Naz, RK, CRC Press, Florida, 293-313 (1993)
14. Tung, K. S. K., Y. Lou and H. Bagavant: Zona pellucida chimeric peptide vaccine. In: *Reproductive Immunology* Eds: Gupta, SK, Narosa Publishing House, New Delhi, India, 303-308 (1999)
15. Naz, R. K. and X. Zhu: Recombinant fertilization antigen-1 causes a contraceptive effect in actively immunized mice. *Biol Reprod* 59, 1095-1100 (1998)
16. Hardy, C. M., H. G. Clarke, B. Nixon, J. A. Grigg, L. A. Hinds and M. K. Holland: Examination of the immunocontraceptive potential of recombinant rabbit fertilin subunits in rabbit. *Biol Reprod* 57, 879-886 (1997)
17. Bussemakers, M. J., A. van Bokhoven, G. W. Verhaegh, F. P. Smit, H. F. Karthaus, J. A. Schalken, F. M. Debruyne, N. Ru and W. B. Isaacs: DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res* 59, 5975-5979 (1999)
18. Liang, P. and A. B. Pardee: Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257, 967-971 (1992)
19. Oettgen, P., E. Finger, Z. Sun, Y. Akbarali, U. Thamrongsak, J. Boltax, F. Grall, A. Dube, A. Weiss, L. Brown, G. Quinn, K. Kas, G. Endress, C. Kunsch and T. A. Libermann: PDEF, a novel prostate epithelium-specific ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression. *J Biol Chem* 275, 1216-1225 (2000)
20. Srikantan, V., Z. Zou, G. Petrovics, L. Xu, M. Augustus, L. Davis, J. R. Livezey, T. Connell, I. A. Sesterhenn, K. Yoshino, G. S. Buzard, F. K. Mostofi, D. G. McLeod, J. W. Moul and S. Srivastava: PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 97, 12216-12221 (2000)
21. Santhanam, R. and R. K. Naz: Novel human testis-specific cDNA: molecular cloning, expression and immunobiological effects of the recombinant protein. *Mol Reprod Dev* 60, 1-12 (2001)
22. Naz, R. K., R. Santhanam and N. Tyagi: Novel human prostate-specific cDNA: molecular cloning, expression, and immunobiology of the recombinant protein, *Biochem Biophys Res Commun* 297, 1075-1084 (2002)
23. Herness, E. A. and R. K. Naz: A novel human prostate-specific gene-1 (HPG-1): molecular cloning, sequencing, and its potential involvement in prostate carcinogenesis. *Cancer Res* 63, 329-336 (2003)
24. Sanger, F., S. Nicklen and A. R. Coulson: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74, 5463-5467 (1977)
25. Kyte, J. and R. F. Doolittle: A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157, 105-132 (1982)
26. Engelman, D. M., T. A. Steitz and A. Goldman: Identifying nonpolar transbilayer helices in amino acid

Human testis, ovary and placenta specific cDNA

sequences of membrane proteins. *Annu Rev Biophys Biophys Chem* 15, 321-353 (1986)

27. Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman: Basic local alignment search tool. *J Mol Biol* 215, 403-410 (1990)

28. Ponte, P., S. Y. Ng, J. Engel, P. Gunning and L. Kedes: Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. *Nucleic Acid Res* 12, 1687-1696 (1984)

29. Naz, R. K: Involvement of protein serine and threonine phosphorylation in human sperm capacitation. *Biol Reprod* 60, 1402-1409 (1999)

30. El-Alfy, M., D. Moshonas, C. R. Morales and R. Oko: Molecular cloning and developmental expression of the major fibrous sheath protein (FS 75) of rat sperm. *J Androl* 20, 307-318 (1999)

31. von Heijne, G.: The signal peptide. *J Membr Biol* 115, 195-201 (1990)

Key Words: Testis/ovary- specific genes, Tissue-specific antigens, Sperm antigens, Contraceptive vaccine, Infertility

Send correspondence to: Dr Rajesh K. Naz, Division of Research, Department of Obstetrics and Gynecology, Health education Building, Rm. 211, Medical College of Ohio, 3055 Arlington Avenue, Toledo, OH 43614-5806, Tel: 419-383-3502, Fax: 419-383-4473, E-mail: Rnaz@mco.edu