1220–1224 *Nucleic Acids Research, 1996, Vol. 24, No. 7*

A nucleolar RNA helicase recognized by autoimmune antibodies from a patient with watermelon stomach disease

Benigno C. Valdez*, Dale Henning, Rose K. Busch, Karen Woods[1], Hernan Flores-Rozas[2], Jerard Hurwitz[2], Laszlo Perlaky and Harris Busch

Departments of Pharmacology and [1]Medicine, Baylor College of Medicine, Houston, TX 77030, USA and [2]Graduate Program in Molecular Biology, Memorial Sloan-Kettering Cancer Center, Sloan-Kettering Institute, New York, NY 10021, USA

Received January 11, 1996; Accepted February 22, 1996

GenBank accession no. U41387

ABSTRACT

Watermelon stomach is characterized by prominent stripes of ectatic vascular tissue in the stomach similar to stripes on a watermelon; in patients with this disorder chronic gastrointestinal bleeding occurs and approximately half of these patients have associated autoimmune disorders. In the serum of one patient, an antinucleolar antibody titer of 1:25 600 was found; the antibodies specifically recognized a ~100 kDa nucleolar protein, which we referred to as the 'Gu' protein. Its cDNA was cloned and sequenced. The Gu protein is a member of a new subgroup of RNA helicases, the DEXD box family. Gu protein fused with glutathione S-transferase contains ATP-dependent RNA helicase activity which preferably translocates in the 5′→3′ direction. Its RNA folding activity, RNA-dependent ATPase and dATPase activities, and its translocation direction are similar to those of RNA helicase II [Flores-Rozas,H. and Hurwitz,J. (1993) *J. Biol. Chem.* 268, 21372–21383]. Sequencing of 209 amino acids of RNA helicase II peptides showed 96.7% identity with the cDNA-derived amino acid sequence of the Gu protein. The precise biological roles of this RNA helicase in the biogenesis of ribosomal RNA and the pathogenesis of watermelon disease and autoimmune disorder require further study.

INTRODUCTION

Watermelon stomach is another term for gastric antral vascular ectasia (GAVE) disease which was first described in 1984. This disease is characterized by prominent stripes of ectatic vascular tissue radiating outward from the pylorus in a form resembling the stripes on a watermelon (1). This disease commonly occurs in older women and may be associated with gastric hemorrhage, iron-deficiency anemia and autoimmune disorders in ~60% of the patients (2,3). The pathogenesis of this disorder and its relationships to autoimmune disease are being studied.

In patient 'Gu', a 1:25 600 titer of antinucleolar antibodies was found. To understand the association of these antibodies to this disease, the cDNA that codes for the autoantigen was cloned. Immunofluorescent staining showed that antibodies produced prominent nucleolar fluorescence. Western blot analysis using this serum stained an ~100 kDa protein, referred to as the Gu antigen, derived from the name of the patient.

Cloning and sequencing of the cDNA for Gu protein is reported in this paper. Based on the cDNA-derived amino acid sequence and the RNA helicase activity of the expressed protein, the Gu protein belongs to a new DEXD box family of RNA helicases (4). These RNA helicases have been implicated in pre-mRNA splicing, translation, ribosomal processing, cell growth and development (5). The cDNA and the derived amino acid sequence of the Gu protein may aid in understanding the relationship of the pathogenesis of this autoimmune disorder associated with watermelon stomach as well as the role of RNA helicases in the biogenesis of ribosomal RNA.

MATERIALS AND METHODS

Cells and serum

HeLa cells were grown at 37°C in Dulbecco's modified Eagle's medium with 10% newborn calf serum and 5% CO_2.

The serum, which was obtained from a patient with watermelon stomach disease, was used in various experiments either as such or as IgG after purification on a protein A affinity column according to the supplier (Pierce).

cDNA cloning and sequencing

Purified immunoglobulins were labeled with Bolton-Hunter Reagent [^{125}I] (ICN) according to the supplier, and the labeled IgGs were used to screen a λgt11 HeLa cDNA expression library (Clontech). A 2 kb insert from positive clone 1A1 was subcloned into Bluescript II KS (Stratagene) and sequenced by the dideoxynucleotide sequencing method using a Sequenase kit (US Biochemical). To find a longer cDNA clone, a 450 bp fragment from the 5′-end of clone 1A1 was ^{32}P-labeled and used to screen the same cDNA library using QuikHyb solution (Stratagene). A positive clone (7A) with a 3.3 kb insert was selected. This λgt11 clone was digested with *Eco*RI and the resulting four insert fragments were subcloned into pTZ18R (Pharmacia Biotech) and

* To whom correspondence should be addressed

both strands were sequenced. The orientation of each fragment was determined by sequencing PCR-amplified products.

Additional 5′ sequence of Gu cDNA was obtained using HeLa poly(A)$^+$ mRNA and 5′ RACE system (Life Technologies). The PCR product obtained using this kit was subcloned at the *Sal*I–*Bam*HI sites of pTZ18R and sequenced as described above.

Western blot analysis and indirect immunofluorescence

HeLa cell nucleolar extract was prepared (6) and immunoblot analyses and indirect immunofluorescence were done as described (7) with 1:500 and 1:10 000 dilutions, respectively, of the patient serum.

Northern blot

Poly(A)$^+$ mRNA was isolated from HeLa cells using an mRNA purification kit (Pharmacia Biotech). Northern blots were done as described (8) using ^{32}P-labeled PCR-amplified cDNA fragments and QuikHyb solution (Stratagene).

Expression and purification of GST–Gu fusion protein

The coding sequence for Gu protein was amplified by PCR and subcloned into the *Bam*HI–*Xho*I sites of pGEX 4T-3 vector (Pharmacia) in frame with glutathione *S*-transferase (GST) gene. The fusion protein was expressed in DH5α *Escherichia coli* cells (Life Technologies, Inc.) and purified from 3 l Luria broth medium using glutathione Sepharose 4B resin as described (9). The protein was eluted from the resin with 50 mM Tris–HCl, pH 8.0, containing 5 mM reduced glutathione and dialyzed overnight against Buffer A (20 mM HEPES–KOH pH 7.6, 0.1 mM EDTA, 1 mM DTT, 10% glycerol) with 0.1 M KCl. The dialyzed fraction was loaded onto a pre-equilibrated 5 ml HiTrap heparin sepharose column (Pharmacia) which was then washed with 50 ml 0.1 M KCl in buffer A. Bound proteins were eluted with a linear gradient of 0.1–1.0 M KCl in buffer A. The GST–Gu fusion protein eluted at ~0.6 M KCl and was visualized by silver staining (BioRad kit). Fractions containing the GST–Gu fusion protein were pooled, concentrated to 0.5 ml using Centricon 30 (Amicon) and desalted by washing two times with buffer A (containing 0.1 M KCl) followed by further concentration using Nanosep 30 (Pall Filtron). The final fraction (80 µl) was aliquoted into 20 µl fractions (1 µg/µl protein) and stored at –80°C until used. Protein concentrations were determined using the BioRad protein assay.

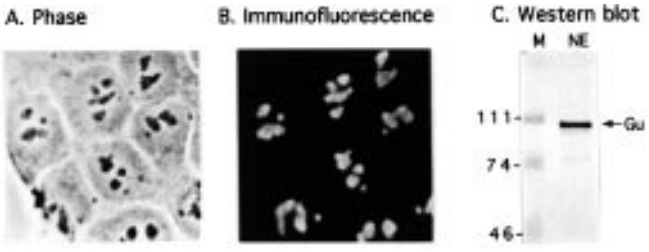

Figure 1. Cellular localization of the Gu protein. Phase (**A**) and indirect immunofluorescence microscopy (**B**) of fixed HeLa cells using watermelon stomach patient serum diluted to 1:10 000 and FITC-coupled antihuman IgG (1:20) taken at a magnification of ×375. The apparent molecular mass of the autoantigen was determined to be 100 kDa by Western blot analysis of nucleolar extract (NE in C). Numbers on the left are molecular mass markers (M) in kDa.

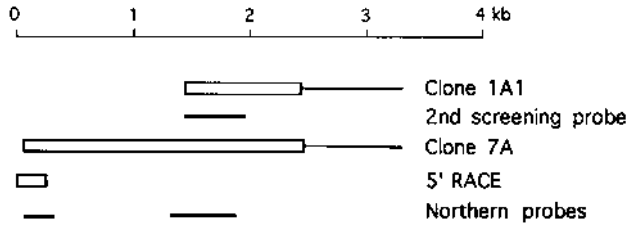

Figure 2. Cloning strategy for the Gu cDNA. HeLa cDNA library screening resulted in identification of clones 1A1 and 7A. Additional 5′ sequence was obtained with a 5′ RACE kit. Open boxes correspond to coding region and thin lines represent the 3′ untranslated region. Thick lines show the positions of the cDNA fragments used for subsequent screening and Northern blot analyses.

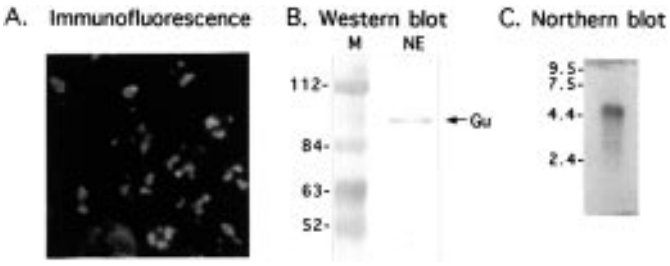

Figure 3. Immunofluorescence microscopy of HeLa cells (**A**) and Western blot analysis of HeLa nucleolar extract (NE in **B**) using undiluted eluted antibodies which bound to Gu protein (×375). The numbers in (B) are molecular mass markers (M) in kDa. (**C**) Determination of the size of the Gu mRNA using the probes shown in Figure 2. The numbers are molecular size markers in kilobases.

Purification of specific anti-Gu antibodies

To obtain antibodies specific for the Gu protein, a blot-affinity purification (BAP) method (10) was used. Glutathione sepharose-purified GST–Gu fusion protein, after electrophoresis through a 10% polyacrylamide–SDS gel, was transferred to a nitrocellulose filter. The position of the GST–Gu fusion protein was identified using strips from the sides of the blot that were stained with amido black. The excised unstained center area of the blot was blocked with 3% milk and then subjected to Western blot. Bound antibodies were eluted with glycine buffer (0.1 M glycine, pH 2.8, 0.5 M NaCl, 0.05% Tween 20), dialyzed against PBS (8.45 mM Na_2HPO_4, 1.55 mM NaH_2PO_4, 0.15 M NaCl, pH 7.5), and then used in immunofluorescence and Western blot analyses.

RNA helicase and ATPase assays

RNA helicase activity was determined using 5′- and 3′-tailed substrates synthesized using Ambion's MAXIscript kit. The 5′-tailed substrate was prepared by transcribing with T3 or T7 RNA polymerase the Bluescript II KS plasmid (Stratagene) cut with *Bam*HI or *Hin*dIII, respectively. The lower strand T7 transcript was labeled with [α-^{32}P]GTP. The preparation of the 3′-tailed substrate, gel purification of all the substrates, RNA helicase assay and ATPase assay were done as described (11) with the exception that 10% polyacrylamide–SDS gels were used for RNA helicase assays.

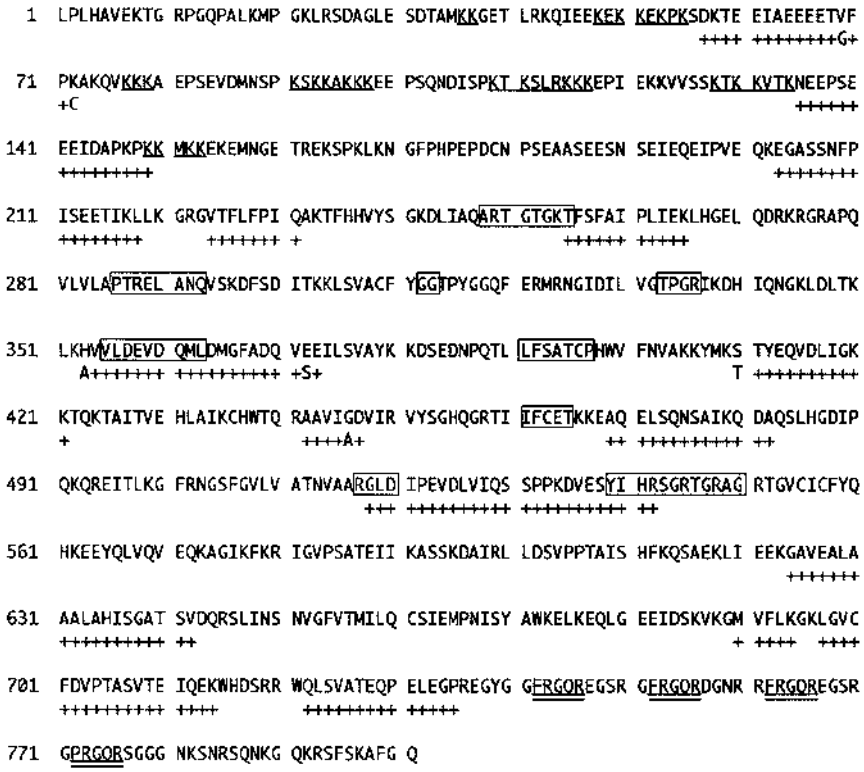

Figure 4. cDNA-derived amino acid sequence of Gu protein. The cDNA sequence was deposited in the GenBank with accession number U41387. Regions rich in lysine residues which may be part of a nuclear localization signal are underlined. Highly conserved regions common to RNA helicases are boxed. Repeats with undetermined function are double-underlined. Sequences of the RNA helicase II peptides (+'s and letters) derived by amino acid sequencing as described under Materials and Methods were compared with the cDNA-derived amino acid sequence of the Gu protein. Identical amino acids are indicated (+) below the cDNA-derived Gu sequence; differences are indicated by the letter codes.

Determination of the peptide sequence of RNA helicase II

A total of 25 µg of the Cibacron blue–agarose fraction containing RNA helicase II (11) was subjected to electrophoresis on a 10% polyacrylamide–SDS gel and then electroblotted onto a nitrocellulose filter and detected by reversibly staining with 0.2% solution of Ponceau-S (Sigma). The band containing the protein was excised. The nitrocellulose immobilized protein was subjected to trypsin digestion as described (12). The resulting peptides were purified through a microbore reverse-phase high performance liquid chromatography (13) and subjected to N-terminal micro-sequencing by Edman degradation (14).

Computer analysis

Nucleic acid and amino acid sequence analyses were done using the Wisconsin Sequence Analysis Package GCG Version 8.

RESULTS AND DISCUSSION

Localization of the Gu protein in the nucleolus

Indirect immunofluorescence of fixed HeLa cells using serum from a patient with watermelon stomach disease revealed the nucleolar localization of Gu protein in HeLa cells (Fig. 1A and B). Staining of MCF7, COS-7, CHO-K1 and NIH 3T3 cells showed similar nucleolar localization (data not shown). Western blot analysis showed that the autoimmune antibodies bound to a specific nucleolar protein with a molecular mass of ~100 kDa (Fig. 1C). These results indicate that the Gu autoantigen in this patient is an immunologically conserved 100 kDa nucleolar protein.

Cloning of the Gu protein cDNA

To identify the antigen recognized by the watermelon disease autoimmune serum, a HeLa cDNA expression library in λgt11 was immunoscreened using purified ^{125}I-labeled IgG. A clone with 2 kb insert (clone 1A1 in Fig. 2) was selected. Sequencing showed that the insert contained a poly(A)$^+$ tail at its 3′-end and 1 kb open reading frame from its 5′-end. Subcloning of this cDNA into a bacterial expression vector and subsequent expression produced a 40 kDa peptide which immunoreacted with the patient's serum (data not shown). To isolate the 5′ portion of the cDNA, the same library was screened with a ^{32}P-labeled probe corresponding to the 5′-end of clone 1A1 (Fig. 2). A clone with 3.3 kb insert was obtained, the 3′-end of which contained the sequences found in clone 1A1. An additional 5′ sequence of 12 nucleotides was obtained using a 5′ RACE (rapid amplification of cDNA ends) kit. The sequence for the Gu protein cDNA was deposited in GenBank with accession number U41387.

To show that the isolated cDNA clones code for the Gu protein, clone 1A1 was expressed as a GST fusion protein and blotted onto nitrocellulose paper. Antibodies from the patient serum were bound to the fusion protein and then eluted. The eluted antibodies recognized a ~100 kDa protein localized in the nucleolus (Fig. 3A and B).

Northern blot analysis of poly(A)$^+$ mRNA isolated from HeLa cells showed that the Gu mRNA is ~4.5 kb long (Fig. 3C). The

same sized mRNA was recognized in two independent experiments using different cDNA probes shown in Figure 2. Based on this result and the sequence of the longest available cDNA clone, Gu-mRNA presumably contains 0.8 kb 5′ untranslated region, an ~2.7 kb coding region, and a 1 kb 3′ untranslated region.

Gu amino acid sequence

Figure 4 shows the cDNA-derived amino acid sequence of Gu protein. The first four amino acid residues were obtained from a clone derived using a 5′ RACE kit and the remaining sequences were derived from clones 1A1 and 7A (Fig. 2). The calculated molecular mass of the available cDNA-derived amino acid sequence is 89 kDa.

Gu protein is a basic protein with combined lysine and arginine residues of 17% and a computed pI of 10.1. The N-terminal region is particularly rich in lysine residues. Groups of lysine residues (underlined in Fig. 4) are putative bipartite nuclear localization signals (15). The C-terminal region is rich in arginine and glycine residues, a domain reported to be essential for efficient binding of nucleolar protein C23 to RNA (16). An RNA binding activity of Gu clone 1A1 (C-terminal region) was observed (data not shown). This region of Gu protein also contains three FRGQR repeats and one PRGQR (double-underlined in Fig. 4) of unknown function.

The middle portion of the molecule contains regions highly conserved in RNA helicases (boxed in Fig. 4). Comparison of the Gu protein sequence with the Peptide Sequence Databases shows the presence of nine motifs conserved in RNA helicases in species ranging from virus to human (4,17,18). Two families of RNA helicases have been reported; the DEAD box family and the DEAH/DEXH box family (4). The Gu protein belongs to the DEXD box family, a variation of the DEAD box family. This variation is similar to DEXH of the DEAH group.

Gu protein is an RNA helicase

RNA helicase II has been purified from nuclear extracts of HeLa cells (11). Peptides of RNA helicase II were sequenced and compared to the cDNA-derived amino acid sequence of the Gu protein. Of these, 202 amino acid sequences (209 residues compared) were identical to those derived from the cDNA protein sequence (Fig. 4). These identities include highly conserved helicase domains and regions that vary widely among different helicases which are believed to be responsible for the specific functions intrinsic to the individual protein (4).

Determination of the RNA helicase activity of the GST–Gu fusion protein showed that 5 ng of the fusion protein could unwind 50 fmol double-strand RNA with 5′ overhangs under the assay conditions and it was 20% efficient in unwinding 50 fmol double-strand RNA with 3′ overhangs (Fig. 5A, lanes 3 and 8). With 2.5 ng of the fusion protein 20% displacement was observed using the 5′ substrate but no activity was observed when 3′ substrate was used (data not shown). No RNA helicase activity was observed in the absence of ATP (Fig. 5A, lanes 5 and 9). Comparison of lanes 2 and 3 (Fig. 5A) showed that the RNA helicase product migrated more slowly than the single-strand RNA on 10% polyacrylamide–0.1% SDS gels. Boiling this product prior to loading onto the gel resulted in a migration identical to the single strand RNA (lane 4). These results imply that the Gu enzyme unwinds the double-strand RNAs and, in addition, catalyzes the folding of the radioactive strand. These results are similar to those observed with RNA helicase II, which exhibited unwinding and folding activities (11).

To test the folding activity of GST–Gu protein, the ^{32}P-labeled T7 transcript of the 5′ substrate was boiled and cooled on ice prior to addition of the GST–Gu protein. Using conditions similar to

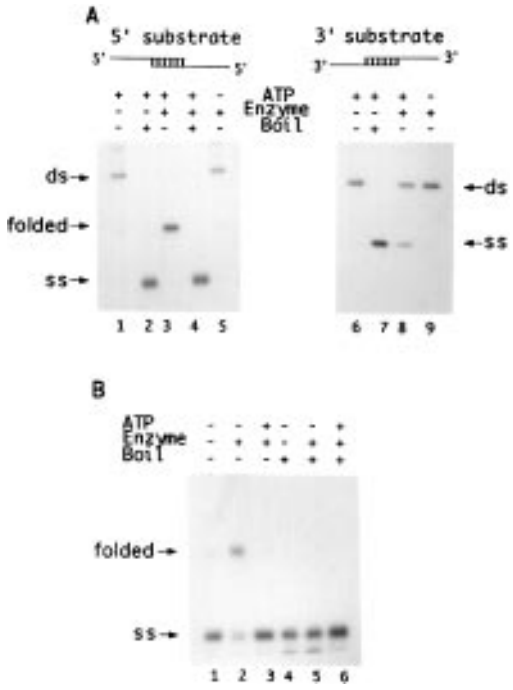

Figure 5. RNA helicase activity of GST–Gu fusion protein. (**A**) RNA helicase activity of 5 ng GST–Gu fusion protein was determined using 5′- and 3′-tailed double-strand RNA substrates as described under Materials and Methods. Reaction products were separated on a 10% polyacrylamide–0.1% SDS gel. Some of the experimental conditions are indicated at the top of each lane. ds, double strand; ss, single strand. (**B**) The ^{32}P-labeled lower strand of the 5′-substrate was boiled for 5 min and cooled on ice for 5 min. This denatured RNA was used as the substrate for the folding activity of 100 ng GST–Gu protein under conditions similar to those used for the RNA helicase assay. Lanes 4–6 are identical to lanes 1–3 but the reaction mixtures were boiled after incubation prior to loading onto a 10% polyacrylamide–SDS gel.

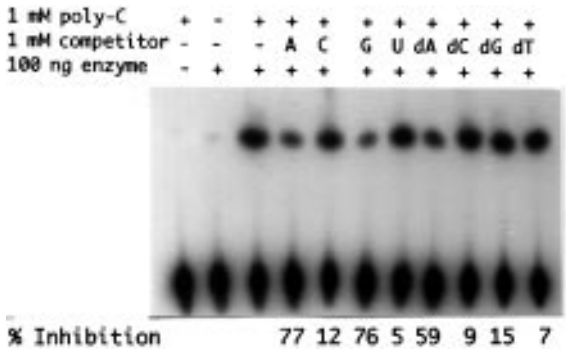

Figure 6. ATPase activity of GST–Gu fusion protein. The ATPase activity of 100 ng GST–Gu fusion protein was determined as described (11) in the absence or presence of unlabeled nucleotides. All components were mixed together except for the enzyme which was added last. Labeled unhydrolyzed ATP (bottom spots) and inorganic phosphate product (top spots) were excised and counted. Inhibition of ATPase activity by the unlabeled nucleotides is indicated by the numbers below the figure. A, ATP; C, CTP; G, GTP; U, UTP; dA, dATP; dC, dCTP; dG, dGTP; dT, dTTP.

the helicase assay, the GST–Gu protein catalyzed the folding of the T7 transcript in the absence of ATP as indicated by a slower migrating band (Fig. 5B, lane 2). In the presence of ATP, no slower migrating band was observed (lane 3). These results indicate that the helicase and RNA folding activities of Gu protein are dependent on the concentration of ATP. The specificity of the folding activity with respect to the length of the RNA and the number of possible hydrogen bonds in the folded structure remains to be determined.

The GST–Gu protein had ATPase activity which was stimulated by poly-C and inhibited by GTP and dATP (Fig. 6). The fusion protein did not contain GTPase activity and its dATPase activity was less than its ATPase activity (data not shown). These results are similar to the NTPase activities of RNA helicase II (11).

The Gu protein has a number of properties similar to RNA helicase II including molecular mass ~100 kDa, ATP-dependent unwinding activity in the 5′→3′ direction, RNA folding activity in the absence of ATP, RNA-dependent ATPase and dATPase activities, and 96.7% identity in the 209 compared amino acid residues. Two other RNA helicases have been reported to be transiently localized in nucleoli and their localization has been reported to be stage dependent. RNA helicase p68 is a nuclear protein which translocates to pre-nucleolar bodies during telophase (19). An3 protein, encoded by a maternal mRNA, is an RNA helicase which co-localizes with the nucleolus only during the mid-vitellogenic stage of *Xenopus* oogenesis (20). These observations imply that p68 and An3 proteins are involved in the early stage of ribosomal RNA biogenesis. The Gu protein has 50% peptide sequence similarity (27% identity) with putative RNA helicases DRS1 and SPB4 yeast proteins which have been shown to be involved in the processing of yeast 27S to mature 25S rRNA (21,22). The role of the Gu protein in ribosomal RNA processing will be studied further.

ACKNOWLEDGEMENTS

This work was supported by the DeBakey and Busch Funds (to H.B.) and the American Heart Association Grant-in-Aid (to J.H.).

REFERENCES

1 Jabbari,M., Cherry,R., Lough,J.O., Daly,D.S., Kinnear,D.G. and Goresky,C.A. (1984) *Gastroenterol.*, **87**, 1165–1170.
2 Gostout,C.J., Viggiano,T.R., Ahlquist,D.A., Wang,K.K., Larson,M.V. and Balm,R. (1992) *J. Clin. Gastroenterol.*, **15**, 256–263.
3 Liberski,S.M., McGarrity,T.J., Hartle,R.J., Varano,V. and Reynolds,D. (1994) *Gastro. Endos.*, **40**, 584–587.
4 Gorbalenya,A.E., Koonin,E.V., Donchenko,A.P. and Blinov,V.M. (1989) *Nucleic Acids Res.*, **17**, 4713–4729.
5 Schmid,S.R. and Linder,P. (1992) *Mol. Microbiol.*, **6**, 283–291.
6 Busch,H., Gyorkey,F., Busch,R.K., Davis,F.M., Gyorkey,P. and Smetana,K. (1979) *Cancer Res.*, **39**, 3024–3030.
7 Perlaky,L., Valdez,B., Busch,R.K., Larson,R.G., Jhiang,S.M., Zhang,W.W., Brattain,M. and Busch,H. (1992) *Cancer Res.*, **52**,428–436.
8 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
9 Valdez,B.C., Perlaky,L., Henning,D., Saijo,Y., Chan,P.K. and Busch,H. (1994) *J. Biol. Chem.*, **269**, 23776–23783.
10 Tang,W.J.Y. (1993) In Asai,D.J. (ed.), *Methods in Cell Biology.* Academic Press, Inc., NY, Vol. 37, pp. 95–104.
11 Flores-Rozas,H. and Hurwitz,J. (1993) *J. Biol. Chem.*, **268**,21372–21383.
12 Tempst,P., Link,A.J., Fleming,M. and Elicone,C. (1990) *Electrophoresis*, **11**, 537–553.
13 Elicone,C., Lui,M., Geromanos,S., Choreda,A. and Tempst,P. (1994) *J. Chrom.*, **676**, 121–137.
14 Erdjument-Bromage,H., Geromanos,S., Choreda,A. and Tempst,P. (1993) In Angeletti,R.H. (ed.), *Techniques in Protein Chemistry IV.* Academic Press, Inc., NY, pp. 419–426.
15 Richter,J.D. and Standiford,D. (1992) In Feldherr,C.M. (ed.), *Nuclear Trafficking.* Academic Press, Inc., NY, pp. 89–119.
16 Ghisolfi,L., Kharrat,A., Joseph,G., Amalric,F. and Erard,M. (1992) *Eur. J. Biochem.*, **209**, 541–548.
17 Wassarman,D.A. and Steitz,J.A. (1991) *Nature* (London), **349**, 463–464.
18 Pause,A., Methot,N. and Sonenberg,N. (1993) *Mol. Cell. Biol.*, **13**, 6789–6798.
19 Iggo,R.D., Jamieson,D.J., MacNeill,S.A., Southgate,J., McPheat,J. and Lane,D.P. (1991) *Mol. Cell. Biol.*, **11**, 1326–1333.
20 Gururajan,R., Mathews,L., Longo,F.J. and Weeks,D.L. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 2056–2060.
21 Sachs,A.B. and Davis,R.W. (1990) *Science*, **247**, 1077–1079.
22 Ripmaster,T.L., Vaughn,G.P. and Woolford,J.L. Jr. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 11131–11135.