Thomas N. Asquith[1]
Dorothy L. Gauggel[1]
Ricardo Esquer-Blasco[2]
N. Leigh Anderson[2]
Robert J. Isfort[1]

[1]Miami Valley Laboratories, The Procter & Gamble Company, Cincinnati, OH, USA
[2]The Large Scale Biology Corporation, Rockville, MD, USA

Confirmed identies of proteins from a two-dimensional map of Syrian hamster embryo cells

The Syrian hamster embryo (SHE) cell transformation assay is widely used to screen chemicals for carcinogenic potential. However, the biochemical mechanisms of transformation in SHE cells are incompletely understood relative to other rodent systems. Thus identification of proteins which change during transformation can provide clues to biochemical mechanisms. Previously, we published a map of SHE cell proteins based on comparisons to other maps. In this report we provide direct sequence analysis of numerous proteins which were previously identified solely by electrophoretic mobility. Protein sequencing verified original spot identifications and extended the range of identified proteins. The updated map will assist in evaluating biochemical mechanisms of morphological transformation in hamster cells.

Keywords: Protein sequence / Hamster / Two-dimensional polyacrylamide gel electrophoresis
EL 3414

Introduction

2-D PAGE is the only method which can simultaneously detect ("map") thousands of proteins in a single gel. However, identifying the proteins remains time-consuming and costly because rigorous identifications of key proteins require extensive biochemical analyses. One alternative is to use master maps as references to make tentative identifications [1, 2] in less characterized maps. Computer analyses are a fast way to tentatively identify proteins. The Syrian hamster embryo (SHE) cell transformation assay is used to study chemical carcinogenesis *in vitro* and as a screen to assess the carcinogenic potential of chemicals. There is an 80–90% concordance between the potential for a chemical to morphologically transform SHE cells and its carcinogenic activity *in vivo* [3, 4]. Changes in cell morphology, growth characteristics, and gene expression are key endpoints of the assay [5, 6]. Identifying proteins that change during transformation will better define the biochemical mechanisms of these endpoints. The hamster proteome is much less characterized than those of rat and mouse. Fortunately, 2-D maps of SHE cells were sufficiently similar to maps from rats and mice that some protein identifications could be extrapolated onto the SHE cell map [7]. The map was developed as a reference to assess changes in protein composition of SHE cells during morphological transformation. Positive protein identifications were needed to verify the tentative assignments and to extend the map to other proteins. Therefore, selected proteins in the hamster map were characterized by protein sequence analysis.

Correspondence: Dr. Tom Asquith, The Procter & Gamble Company, Miami Valley Laboratories, P.O. Box 538707, Cincinnati, OH 45253-8707, USA
E-mail: asquithtn@pg.com
Fax: +513-627-1940

Abbreviation: SHE, Syrian hamster embryo

2 Materials and methods

2.1 SHE cell culture conditions

Cells were grown and proteins were collected as previously described [7].

2.2 2-D PAGE of gels for protein analyses

Analyses were done with an ISO-DALT system (Hoefer, San Francisco, CA) essentially as recommended [8]. Briefly, cells were lysed in 9 M urea / 4% w/v CHAPS / 2% v/v pH 9–11 carrier ampholytes / 1% w/v DTT and clarified by centrifugation. Isoelectric focusing was done in tube gels (14 × 0.1 cm) of 5.7% w/v acrylamide / 0.15% w/v piperazine diacrylamide. The gels contained 9 M urea / 2% v/v NP-40 / 1.4% v/v pH 3–10 carrier ampholytes / 0.6% v/v pH 5–8 carrier ampholytes. The gels were prefocused before loading the samples. Carbamoylated p*I* standards (Pharmacia, Piscataway, NJ) were run as internal standards with selected samples. Gels were electrophoresed overnight for a total of about 13 000 Vh with 20 mM sodium hydroxide and 0.85% phosphoric acid as the cathode and anode buffers, respectively. Immediately after IEF, the tube gels were extruded onto a loading platform, briefly incubated with Laemmli running buffer and then loaded onto the slab gels (16 × 16 × 0.1, 8–18% gradient of cross-linker with 0.1% SDS / 375 mM Tris-HCl; Integrated Separation Systems, Natrick, MA). IEF gels and a separate gel of M_w standards were sealed against the slab gels with 3–4 drops of hot, 0.5% agarose solution

 0173-0835/99/0707-1646 $17.50+.50/0

(24 mM Tris/200 mM glycine / 0.1% w/v SDS / 0.5% w/v agarose). Slab gels were run overnight in standard Laemmli running buffer cooled to about 4°C at constant voltage.

2.3 Electroblotting and staining conditions

Proteins were electrophoresed at constant current (150 mA/gel) for about 2 h onto Problott PVDF (Applied Biosystems, Foster City, CA). The running buffer was preferably 20% v/v MeOH / 25 mM Tris / 192 mM glycine or 10 mM CAPS (pH 9.5). Blots were extensively rinsed with MQ water to remove traces of buffer salts. Dried blots were wetted with MeOH and pretreated (5% w/v trichloroacetic acid / 10% v/v glacial acetic acid / 10% v/v MeOH) for about 3 min to acidify the proteins. Blots were rinsed with 2 × 100 mL 25% MeOH and stained with 0.05% Coomassie Brilliant Blue R-250 for about 4 min. Blots were destained with 3–4 × 100 mL rinses of 60% v/v MeOH / 10% v/v glacial acetic acid, followed by multiple rinses with MQ water. Blots were stored in sealed plastic bags at 0°C between sheets of filter paper.

2.4 Staining of gels with colloidal Coomassie Blue G

The procedure was according to the modified version of Neuhoff *et al.* [9]. Briefly, gel were washed with MQ water, fixed with 10% v/v MeOH / 10% w/v TCA / 10% v/v acetic acid, washed again with MQ water and soaked overnight in 1.6% v/v phosphoric acid/12% w/v ammonium sulfate / 20% v/v MeOH. Gels were then soaked for about 48 h in colloidal Coomassie Blue G (13.8% w/v ammonium sulfate / 1.7% v/v phosphoric acid / 0.085% w/v Coomassie Blue G). Destaining was done by rocking gels in 1.6% v/v phosphoric acid / 12% w/v ammonium sulfate / 20% v/v MeOH.

2.5 Tryptic digestions

Blocked proteins were excised from gels for tryptic digests and HPLC mapping. Gel slices were extensively rinsed with MQ water followed by digestion buffer. Washed gel slices were crushed in digestion buffer (1 M urea / 50 mM Tris-HCl / 172 mM NaCl / 0.05% Triton X-100 / 0.25 mM EDTA) *via* Teflon pestle within 1.5 mL centrifuge tubes. Cysteines were reduced in 2.5 mM DTT and then alkylated in 5 mM *N*-isopropyliodoacetamide (NIPIA; Molecular Probes, Portland, OR) for about 30 min at room temperature in the dark. Sequencing grade trypsin (Boehringer-Mannheim, Indianapolis, IN) was added to each gel solution. All samples were initially placed at room temperature for 4 h followed by overnight incubation at 30°C.

2.6 Chromatography of peptides

Gel fragments were removed *via* filtration through a 0.22 micron filter (Millipore, Bedford, MA) and centrifugation (13 000 rpm). Peptides were concentrated by injecting multiple aliquots onto the column prior to the start of chromatography. Digests were analyzed with a Waters 625 (Milford, MA) pump system, 600E controller and 484 detector. The column was a Waters Delta-Pak C18, 300 Å, 2 mm × 15 cm. Mobile phases were 0.1% TFA in Milli-Q water and in acetonitrile. The flow rate was 0.25 mL/min. Absorbance was monitored at 214 nm. Data were collected with a 900 Nelson databox and processed *via* Turbochrome software. Separations were done with a linear gradient (98:2–40:60) followed by isocratic conditions (40:60) from 40 to 50 min.

2.7 Protein sequencing

Amino acid sequences were determined by Edman degradation. Proteins were excised from blots or peptides were immobilized on Biobrene-treated filters. Analyses were done with an Applied Biosystems 476A model sequencer and data were compiled with the 610 software. Standard cycles (FSTNML) and HPLC conditions were used. Sequences were searched against databases using ENTREZ.

2.8 Scanning and image analysis for master gel pattern

Each stained slab gel was digitized in red light at 133 micron resolution using an Eikonix 1412 scanner. Raw digitized gel images were archived on optical disks, and a greyscale videoprint was prepared from the raw digital image as a hard-copy backup of the gel image. Each 2-D gel was processed using the Kepler software system with procedure PROC008b to yield a spotlist giving position, shape, and density information for each detected spot. This procedure makes use of digital filtering, mathematical morphology techniques, and digital masking to remove background and uses full two-dimensional least-squares optimization to refine the parameters of a 2-D Gaussian shape for each spot. Processing parameters and file locations were stored in a relational database, while various log files detailing operation of the automatic analysis software were archived with the reduced data. The computed resolution and level of Gaussian convergence of each gel were inspected and archived for quality control purposes.

2.9 Image mapping

The SHE cell 2-D pattern was compared to reference 2-D maps obtained using the same procedures described in

this paper from the well-characterized human, rat and mouse liver systems. Proteins with highly conserved amino acid sequences among different mammalian species should yield similar chemical properties, including M_w and p*I* resulting in similar inter-species electrophoretic mobility in both dimensions as well as spot shape. One can also expect similar abundances for these proteins, especially between closely related species. Consequently, a set of previously identified spots, corresponding to such proteins and showing analogous position, shape and abundance in the reference maps, were tentatively identified in the SHE pattern [7]. A processed, scanned SHE cell gel image was generated in Tag Image File Format (TIFF) to create the map in Fig. 1. The notations were manually added using the PC-based presentation software Lotus Freelance 2.1.

3 Results and discussion

The updated map is shown in Fig. 1. Table 1, which lists map labels of the proteins, includes proteins identified either by amino acid sequencing or by electrophoretic mobility. A total of eighteen proteins was sequenced. The identities of β-actin (ACTB), collagen-like protein (CLP), heat-shock protein 60 (HSP 60), heat-shock protein 70 (HSP 70), protein disulfide isomerase (PDI), β-tubulin (TBB) and vimentin (VIME) in the original map [7] were confirmed by protein sequencing. Calreticulin (CRTC),

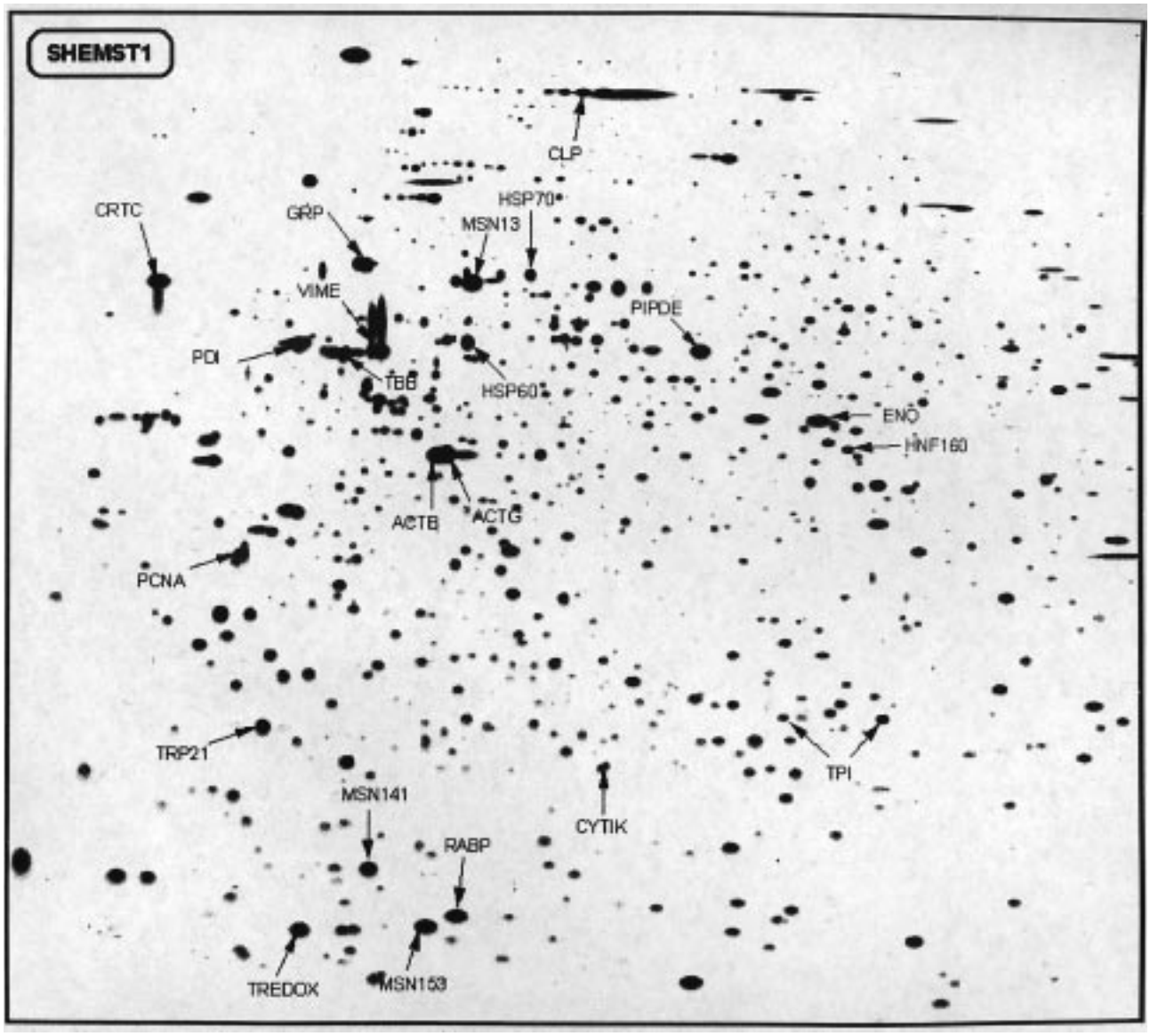

Figure 1. Master 2-D gel pattern of Syrian hamster embryo proteins with protein identifications. Spot numbers are detailed in Table 1. The master image and spot numbers are available from Large Scale Biology.

Table 1. SHE cell proteins identified by amino acid sequencing

Protein	Map label	Map number
β-Actin	ACTB	45
γ-Actin	ACTG	48
α-Type (1) collagen	CLP	8, 11, 24, 69
Calreticulin	CRTC	5, 12
Cytidylate kinase	CYTIK	60 (right) or 547 (left)
Enolase	ENO	85
78 kD Glucose-regulated protein	GRP78	10, 15, 364
Unknown, homology to 160 kD neurofilament	HNF 160	175
60 kD Heat shock protein	HSP60	27
70 kD Heat shock protein	HSP70	22
Proliferating cell nuclear antigen	PCNA	95 or 183, 245 or 317
Protein disulfide isomerase	PDI	17, 36, 79
Phosphatidylinositol-4,5-biphosphate phosphodiesterase	PIPDE	43
Retinoic acid-binding protein	RABP	161
β-Tubulin	TBB	29, 92
Triose phosphate isomerase	TPI	277, 392
Thioredoxin	TREDOX	144
21 kD Tumor-regulated protein	TRP21	114
Vimentins	VIME	14, 29, 23, 33

Map numbers refer to the master map in Isfort *et al.* [6, 7].

enolase (ENO), 78 kDa glucose-regulated protein (GRP), proliferating cell nuclear antigen (PCNA), phosphatidyl-4,5-biphosphate phosphodiesterase (PIPDE), retinoic acid-binding protein (RABP), triose phosphate isomerase (TPI), thioredoxin (TREDOX) and 21 kDa tumor-regulated protein (TRP) were identified by sequencing and assigned to spots in the original map. Of these, TPI was sequenced from two different spots. CYTIK and PCNA were assigned to small regions of the map due to ambiguous spot assignments between the blots and the master map. HNF 160 was an unknown protein with homology to neurofilament protein 160. Gamma-actin (ACTG) was identified by mapping comparison based upon M_w, p*I*, and relative abundance. Map numbers are included for image matching *via* the master map (Large Scale Biology, Rockville, MD).

There is relatively little published information for hamster sequences and even less is available on amino acid sequences for hamster proteins. Amino acid sequence was available for the glucose-regulated protein (A. S. Lee/Gene Bank) and DNA sequences were available for ACTB, GRP78, HSP 60, TBB and VIME. Therefore identifications were done against sequences from mouse and rat. As expected, most hamster proteins were highly homologous to rat and mouse proteins. These data were consistent with Nakatani *et al.* [10] who suggested that hamsters are a separate order within Rodentia. CYTIK was compared against pig (*Sus scrofa*) because there were no published sequences from rodents. The larger variations for CLP were probably caused by comparison of related proteins or by incomplete sequence from peptides rich in proline and glycine. Table 2 lists *N*-terminal amino acid sequences of hamster proteins for which there was no published sequence information. Table 3 lists selected internal amino acid sequences of hamster proteins for which there was no published sequence information and which had blocked *N*-termini. Other internal sequences for CLP, PDI, ENO, ACTB, CYTIK, PIPDE and VIME are not shown. Note that the amino acid sequences for ACTB, GRP78, HSP60, TBB and VIME were consistent with published hamster DNA sequences (data not shown).

In many cases one identity was assigned to multiple spots in the original map. Groupings were made based on known patterns for a given protein in other animals. Often a protein has multiple forms due to charge heterogeneity, glycosylation and protease processing. Identification of one spot strengthened the proposed identities of related spots. For example, spots 14, 19, 23 and 33 in the original map [7] were proposed to be vimentin. Identification of spot No. 14 as vimentin made identities of 19, 23 and 33 more likely to be vimentin. Identification of new proteins for the map allowed us to assign identities to additional spots based on the known heterogeneity of these proteins in other 2-D PAGE patterns.

2-D PAGE remains the best method for simultaneous analysis of thousands of proteins. However, little useful information about the biology of the system can be obtained unless the identities of relevant proteins (spots)

Table 2. *N*-Terminal protein sequences of hamster and rodents

Protein	Animal	Sequences	References	Identity[a)]
CYTIK	Hamster	MKPLV VFVLQ xPGAg KQxQQ AQLV		15/22
	Pig	MRPKV VFVLG GPGAG KGTQC ARIV	[13]	
HSP70	Hamster	AGEAI KGAV vg IDLGt tNg		18/19
	Mouse	ASEAI KGAV VG IDLGT TNG	[14]	
PIPDE	Hamster	SDVLG LTDEN FESR		13/14
	Rat	SDVLE LTDEN FESR	[15]	
PCNA	Hamster	qFEAR LVQGx ILKKV LEAL		16/19
	Mouse	MFEAR LIQGS ILKKV LEAL	[16]	
PDI	Hamster	DAPEE EDNVL VLKSN FAEAL		18/20
	Mouse	DALEE EDNVL VLKKS N FEEAL	[17]	
RABP	Hamster	PNFAG TwKmR qSENFD		14/15
	Mouse	PNFAG TWKMR SSENFD	[18]	
TREDOX	Hamster	VKLIE SKEAF QEAL		13/14
	Rat	M VKLIE SKEAF QEAL	[19]	
TPI	Hamster	a/qPxRK FFVGg/q NxKt		12/12
	Mouse	APTRK FFVGG NWKT	[20]	
TRP21	Hamster	MIIYx DIISH DELFS DIYKIR		19/20
	Mouse	MIIYR DLISH DELFS DIYKIR	[21]	

a) Based only on identified residues. Numbers refer to homolog over total residues.
Lower-case letters are tentative calls and "x" refers to no call.
Two lower-case letters indicate that two residues were detected.

Table 3. Internal protein sequences of hamster and rodents

Protein	Animal	Sequences	References	Identity[a)]
CRTC	Hamster	HEQNI DxGGG YVKL		13/13
	Rat	HEQNI DCGGG YVKL	21	
CLP	Hamster	PICPA PQDNR PGPAG PPGA		15/19
	Mouse	PPGPA GQDGR PGPAG PPGA	22	
ENO	Hamster	xxAR EIIDS RGNPT		11/12
	Mouse	KIFAR EILDS RGNPT	23	
HNF160	Hamster	KVEAi/k KLNVr/l	24	7/10
	Mouse	KVEAP KLKVQ		

a) Based only on identified residues. Numbers refer to homologous residues over total residues.
Lower-case letters are tentative calls and "x" refers to no call. Two lower-case letters indicate that two residues were detected.

are known. One approach is to make tentative assignments in closely related species based upon comparisons against well-characterized maps. For example, Anderson *et al.* [1] tentatively identified numerous proteins in mouse liver based upon comparisons to a 2-D map of rat liver proteins. The original SHE cell map had tentative identities of numerous proteins based upon similarities between the SHE cell pattern and 2-D PAGE patterns from other mammals. Protein sequencing verified the identities of every protein tested. Through image matching it should be possible to assign identities to more hamster proteins from maps of other rodents. The ongoing efforts to identify and characterize rat and mouse proteins should provide substantial information for mapping between species. Similarly, proteins identified in the SHE cell map could be used as references to identify proteins in other maps. Use of immobilized carrier ampholytes should improve matches between gel pattern from different groups. Given enough similarity between two maps, it may be possible to infer if similar cellular responses occur between two systems and so reduce usage of animals.

Definitive identification should be done by chemical analyses because image mapping is based upon electrophoretic mobility. Wilkens and Gooley [11] clearly spell out sources of variation protein mobilities and therefore matching. Computer matches are best for conserved proteins between related species. For example, β- and γ-

actins have highly conserved amino acid sequences and have similar 2-D gel patterns between related organisms. However, similar amino acid sequences do not guarantee similar electrophoretic mobilities. A few changes in amino acid sequence can yield charge shifts from differing post-translational modifications and/or SDS mobility shifts [12].

4 Concluding remarks

Reasonably accurate spot assignments can be made in 2-D maps of less characterized species when compared against well-characterized reference maps from related species. Maps of hamster cells are compatible with maps of rat and mouse cells. By comparing 2-D maps it could be possible to determine if similar biological phenomena occur between a well-characterized model and one which is less characterized. The updated map will assist in evaluating biochemical mechanisms in cultured hamster cells.

We thank Ms. A Fieno for help in sequence searches and Mr. R. Grant for editorial comments.

Received November 12, 1998

5 References

[1] Anderson, N. L., Esquier-Blasco, R., Hofmann, J.-P., Meheus, L., Raymackers, J., Steiner, S., Witzmann, F., Anderson, N. G., *Electrophoresis* 1995, *16*, 1977–1981.

[2] Sanchez, J.-C., Appel, R. D., Golaz, O., Pasquali, C., Ravier, F., Bairoch, A., Hochstrasser, D. F., *Electrophoresis* 1995, *16*, 1131–1151.

[3] Barret, J. C., Hesterberg, T. W., Thomasse, D. G., *Pharm. Rev.* 1984, *36*, 53S–70S.

[4] Oshimura, M., Hesterberg, T. W., Tsutsui, T., Barrett, J. C., *Cancer Res.* 1984, *44*, 5017–5022.

[5] Isfort, R. J., Cody, D. B., Doersen, C. J., Kerckaert, G., Leboeuf, R., *Int. J. Cancer* 1994, *59*, 114–125.

[6] Isfort, R. J., Cody, D. B., Asquith, T. N., Leboeuf, R., *In Vitro Cell. Dev. Biol.* 1992, *28A*, 635–642.

[7] Isfort, R. J., Kerckaert, G., Anderson, N. L., Leboeuf, R., *Electrophoresis* 1992, *13*, 855–861.

[8] Anderson, N. L., *Two-dimensional Electrophoresis-Operation of the ISO-DALT System*, Large Scale Biology Press, Washington DC 1988.

[9] Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W., *Electrophoresis* 1988, *9*, 255–262.

[10] Nakatani, T., Suzuki, Y., Yoshida, K., Sinohara, H., *Biochim. Biophys. Acta* 1995, *1263*, 245–248.

[11] Wilkens, M. R., Gooley, A. A., in: Wilkens, M. R., Williams, K. L., Appel, R. D., Hochstrasser, D. F. (Eds.), *Proteome Research: New Frontiers in Functional Genomics*, Springer-Verlag, Berlin 1997, pp. 35–64.

[12] Anderson, N. L., Anderson, N. G., *Biochem. Biophys. Res. Comm.* 1979, *88*, 258–265.

[13] Okajima, T., Goto, S., Tanizawa, K., Tagaya, M., Fukui, T., Shimofuruya, H., Suzuki, J., *J. Biochem.* 1995, *117*, 980–986.

[14] Giometti, C. S., *Electrophoresis* 1992, *13*, 970–991.

[15] Bennet, C. F., Balcarek, J. M., Varrichio, A., Crooke, S. T., *Nature* 1988, *334*, 268–270.

[16] Yanaguchi, M., Hayashi, Y., Hirose, F., Matsuoka, S., Moriuchi, T., Shiroishi, T., Moriwaki, K., Matsukage. A., *Nucleic Acids Res.* 1991, *19*, 2403–2410.

[17] Mazzarella, R. A., Srinivasan, M., Haugejordan, S. M., Green, M., *J. Biol. Chem.* 1990, *265*, 1094–1101.

[18] Wei, L. N., Tsao, J. L., Chu, Y. S., Jeannotte, L., Nguyen-Huu, M. C., *DNA Cell Biol.* 1990, *9*, 471–478.

[19] Tonissen, K. F., Robins, A. J., Wells, J. R. E., *Nucleic Acids Res.* 1989, *17*, 3973.

[20] Cheng, J., Mielnicki, L. M., Pruitt, S. C., Maquat, L. E., *Nucleic Acids Res.* 1990, *18*, 4261.

[21] Chitpatima, S. T., Makrides, S., Bandyopadhyay, R., Brawerman, G., *Nucleic Acids Res.* 1988, *16*, 2350.

[22] Murthy, K. K., Banville, D., Srikant, C. B., Carrier, F., Holmes, C., Bell, A., Patel, Y. C., *Nucleic Acids Res.* 1990, *13*, 4933.

[23] French, B. T., Lee, W. H., Maul, G. G., *Gene* 1985, *39*, 311–312.

[24] Lamande, N., Brosset, S., Keller, A., Lucas, M., Lazar, M., *EMBL Data Library, 9*, 1991.

[25] Levy, E., Liem, R. K. H., D'Eustachio, P., Cowan, N. J., *Eur. J. Biochem.* 1987, *166*, 71–77.