BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **242,** 79–83 (1998)
ARTICLE NO. RC977920

Oligomerization of N-Terminal Domain of Carcinoembryonic Antigen (CEA) Expressed in *Escherichia coli*

Anna Krop-Watorek,* Shinzo Oikawa,† Yoshiaki Oyama,† and Hiroshi Nakazato†,[1]
**Institute of Immunology and Experimental Therapy, Czerska 12, 53-114 Wroclaw, Poland; and*
†Suntory Institute for Biomedical Research, Osaka 618, Japan

Received November 25, 1997

The N-terminal domain of CEA, which is essential for cell adhesion activity and lacks cysteine residue, was expressed in *Escherichia coli* and purified from the solubilized inclusion bodies by DEAE-Sepharose and gel filtration chromatographies. The purified N-domain migrated in SDS-PAGE as a single 13-kDa band, whereas it migrated in non-SDS-PAGE as five distinct bands. The N-domain, analyzed by two-dimensional PAGE after cross-linking with DSS, migrated in multiple forms ranging from monomer to pentamer, showing unequivocally the presence of multimers in each band. The amount of monomer was distinctively the least among the oligomers in the non-SDS-PAGE. These results suggest that the N-domain of CEA molecule has a strong tendency to self-assemble that may convey the homophilic cell adhesion of CEA.

CEA is one of the best characterized human tumor-associated antigens. The serum level of CEA is frequently used in diagnosis and monitoring of cancer patients following surgery mainly for colon, breast and lung adenocarcinomas (see 1 for review).

The primary structure, deduced from molecular cloning of the CEA (2-4) and its related antigens (5–8) revealed that human CEA family belongs to the immunoglobulin (Ig) superfamily (9, 10) consisting mainly of one to seven Ig-like domains.

CEA is a glycoprotein of a molecular mass of about 180 kDa in SDS-PAGE. It carries seven Ig-like domains connected in tandem, the N-domain being unique in that it lacks Cys residue while each of the other six domains carries two Cys residues that apparently forms a disulfide bridge. The C-terminal, 26-residue hydrophobic domain, is post-translationally replaced by GPI-anchor (11, 12).

The biological functions of CEA family members are largely unknown at present. However, their cell surface localization, structural similarities to the Ig-superfamily and the presence of RGD tripeptide in the amino acid sequences of some members, suggest a role in cell-cell and cell-substrate interactions (9, 13). In the earlier studies it was found that CEA and NCA interacted in the solution forming homodimers and even higher aggregates (14, 15). The rat homologue of BGP, C-CAM, was also shown to exist as non-covalently linked dimers both in solution and on the cell surface (16). Recent findings that CEA, NCA and BGP are able to mediate homotypic and/or heterotypic cell adhesion *in vitro*, suggest that they might also function as cell adhesion molecules *in vivo* (17–24). Furthermore, it was shown that N-domains are essential for the CEA-mediated homophilic as well as heterophilic adhesion between the members of CEA family, NCA and CGM6 (W272) (25).

To elucidate the biological roles of CEA family members, it is important to know the biochemical characteristics of the molecules which requires substantial amounts of them. *E. coli* expression system is suitable for producing large amounts of recombinant proteins, albeit the ease of the production differs from protein to protein. Its products not being glycosylated, it is also useful for the comparative study of the structure-activity relationship of a protein in non-glycosylated and glycosylated forms.

In the previous studies we noticed that N-domains of CEA and NCA produced in *E. coli* could retain conformational epitopes while the other domains were produced but in very different conformations than the native ones (26).

[1] To whom correspondence should be addressed. Fax/: +81-75-962-6448. E-mail: nakazato@minase.SUNTORY.co.jp.

Abbreviations used: BGP, biliary glycoprotein; CAM, cell adhesion molecule; CEA, carcinoembryonic antigen; DSS, disuccinimidyl suberate; GPI, glycosylphosphatidylinositol; NCA, nonspecific cross-reacting antigen; PBS(−), phosphate buffered saline without Mg++ and Ca++; PSG, pregnancy specific β1-glycoprotein; RGD, Arg-Gly-Asp sequence.

0006-291X/98 $25.00

We report here that CEA N-domain expressed in *E. coli* has self-associative properties and forms oligomers composed of up to five units. The role of self-assembling activity of the N-domain in CEA homophilic adhesion is discussed.

MATERIALS AND METHODS

Materials. Silver staining kit was from Daiichi Pure Chemicals Co. (Tokyo, Japan). DEAE-Sepharose, Sephacryl S100 was from Pharmacia LKB Biotechnology Inc. (Uppsala, Sweden). Pre-stained molecular weight markers were from Bio-Rad (Richmond, USA). DSS was from Pierce (Rockford, USA). Urea was of specially prepared reagent grade from Nacalai Tesque (Kyoto, Japan).

Construction of expression vector. The fragment encoding N-domain of CEA was amplified by PCR using a DNA Thermal Cycler (Perkin-Elmer Cetus, CA), in 100 μl of a standard reaction mixture of the Gene-Amp DNA Amplification Reagent Kit (Perkin-Elmer Cetus) following the manufacturer's instruction. The oligonucleotides 5′-TGGAACCCGCCCGAATTCATGAAGCTCACTATT-3′, containing an *Eco*R I site and an initiating methionine codon and 5′-TTTCTA-GATTACGGGTATACCCGGAACT-3′, containing an *Xba* I site and a stop codon were used as the 5′- and 3′-primers, respectively. The amplification reaction was carried out by heating at 94 °C for 1 min, followed by annealing at 60 °C for 1 min and extension at 72 °C for 1 min. This reaction was repeated for 30 cycles followed by a 10 min final extension at 72 °C. The PCR product (300 bp) was purified by 2% agarose gel electrophoresis, digested with *Eco*R I and *Xba* I and ligated into the *Eco*R I and *Xba* I sites of pUC-PL-cI (26) to give pUC-PL-CEA/108. The DNA sequencing of pUC-PL-CEA/108 revealed that there was no base change in the PCR amplified region.

Expression and production of N-domain in E. coli. Transformation of competent *E. coli* W3110 was performed following the manufacturer's instruction. The transformant was subcultured overnight at 32 °C in 2 × 300 ml of super broth (SB) medium containing ampicillin (50 μg/ml). Each of four fermenter tanks containing 3 liters of SB/ampicillin medium was inoculated with 150 ml of overnight culture and cell growth was continued for 3 hrs at 32 °C until the OD660 reached 5.0. Then, temperature was shifted to 42 °C and the culture was continued for 3 more hrs, to the late log phase of growth.

Purification of recombinant N-domain. All steps were performed at 4 °C. The *E. coli* cells were collected by centrifugation for 10 min at 8,000 rpm, resuspended in 2 liters of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), disrupted in a Manton Gaulin Cell Disrupter and centrifuged as above. The precipitate of inclusion bodies was washed consecutively with 1 liter of TE, 500 ml of TE containing 0.5% Triton X-100 and then with 500 ml of TE. The first and the second washing was followed by centrifugation at 8,000 rpm for 30 min, the third one at 12,000 rpm for 10 min. The washed inclusion bodies were frozen and stored at −20 °C until use. The thawed inclusion bodies were solubilized overnight by gently stirring in 50 mM Tris-HCl, pH 8.5, containing 6 M urea, 5 mM DTT and 1 mM EDTA (solubilization buffer). The solution was clarified by centrifugation at 35,000 rpm for 30 min and applied on a DEAE-Sepharose column (1.5 × 20 cm) equilibrated with the same buffer. The unbound fractions, containing the N-domain, were pooled, concentrated in a Centriprep 10 Amicon Concentrator and applied on a Sephacryl S100 column (1.5 × 100 cm) equilibrated with the solubilization buffer, supplemented with 0.5 M NaCl. The fractions after both columns were analyzed by SDS-PAGE and Western blotting. The fractions containing N-domain were pooled and dialyzed against four changes of 2 liters of PBS(−) for at least 48 hrs at 4 °C.

Polyacrylamide gel electrophoreses (PAGE). SDS-PAGE was carried out essentially as described by Laemmli (27). The samples in SDS-containing sample buffer were heated at 100 °C for 5 min, then separated in 15% gel. Non-SDS-PAGE was carried out in 10% gel, the samples were neither treated with SDS nor heated. For the two-dimensional electrophoresis, lanes from the non-SDS gel were cut off, soaked for 30 min in a sample buffer with or without SDS, sealed over the stacking gels for the 2nd direction and electrophoresed at 100 V for 2 hrs. Proteins in the gels were visualized with Coomassie Brilliant Blue R250 or by silver staining.

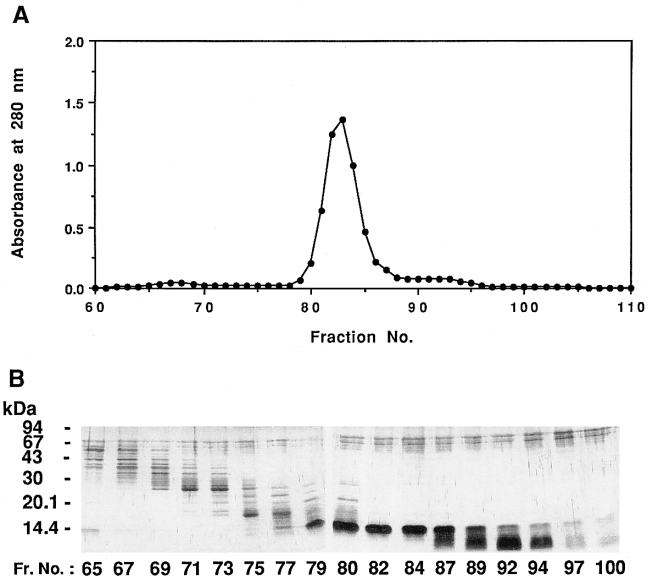

FIG. 1. Sephacryl S100 column chromatography (A) and SDS-PAGE (B) of the recombinant CEA N-domain. In A, 2 ml of 0.4% protein solution obtained from DEAE-Sepharose was loaded on a column (1.5 × 100 cm) equilibrated with 50 mM Tris-HCl, pH 8.5, containing 6 M urea, 5 mM DTT, 1 mM EDTA, and 0.5 M NaCl. Each fraction (0.7 ml) was collected at a flow rate of 2 ml/hr. In a pre-run, BSA was eluted at around fraction 67. Protein concentration was measured by absorbance at 280 nm. In B, 10-μl aliquots from indicated column fractions were taken and analyzed by SDS-PAGE under reducing conditions. Proteins in gel were developed by silver staining. The molecular weight markers, phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20.1 kDa), and α-lactalbumin (14.4 kDa), were used. Fraction numbers and molecular weight markers are shown on the bottom and at the left, respectively.

Cross-linking. DSS at the concentration of 50 mM was freshly dissolved in dimethyl sulfoxide. CEA N-domain (250 μg/ml) in PBS(−) containing 4.5 mM DSS (28) was incubated at 20 °C for 2 hrs. The reaction was stopped by adding 1 M Tris-HCl buffer, pH 7.5 to a final concentration of 80 mM.

RESULTS

Expression and purification of CEA N-domain. Fig. 1 depicts the elution profile through a Sephacryl S100 gel filtration column of the proteins that did not bind to DEAE-Sepharose column. Most of the proteins were eluted in a single major peak. Fractions 82 to 84 which contained purified CEA N-domain that migrated in a single band of *ca.* 13 kDa upon SDS-PAGE analysis (Fig. 1B) were pooled and renatured by dialysation against PBS(−) for 48 hrs (see Materials and Methods).

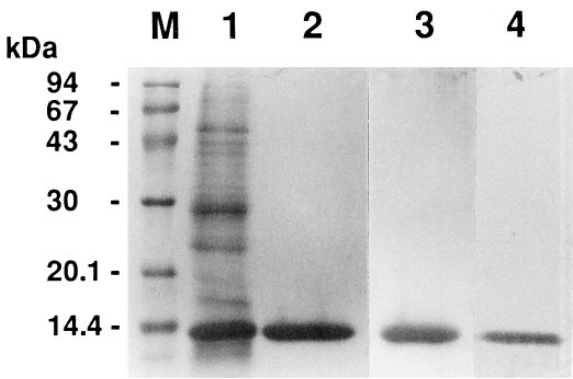

FIG. 2. SDS-PAGE analysis of the recombinant CEA N-domain. SDS-PAGE 15% gel was performed as described in Materials and Methods: lane 1, inclusion bodies before purification; lane 2, purified CEA N-domain before dialysis; lanes 3 and 4, purified CEA N-domain after dialysis against PBS(−). Lanes 1, 2, and 3 were performed under reducing conditions, lane 4 under non-reducing conditions. Lane M, molecular weight markers.

The size of 13 kDa is in good agreement with 12.1 kDa calculated from the deduced amino acid sequence for the non-glycosylated N-domain (Figs. 1B and 2).

Fig. 2 depicts the SDS-PAGE patterns of inclusion bodies before purification (lane 1) and of purified N-domain (lanes 2–4). The N-domain migrated as a single band under both reductive and oxidative conditions (Fig. 2 lanes 3 and 4) reflecting the absence of Cys residues in the sequence.

Characterization of purified CEA N-domain. Upon non-denaturing PAGE, the recombinant N-domain migrated in five distinct bands (Fig. 3B, lane 1) which suggested the presence of either multimeric forms, conformational isomers or charge isomers. The charge isomers could have been produced by carbamylation of Lys residues by exposure to urea during the solubilization of the inclusion bodies and the following purification steps, albeit the precautions taken to minimize the reaction, *e.g.*, use of freshly prepared ultra-pure urea solution, low temperature and minimum exposure to urea.

In order to distinguish between the multimeric forms and monomeric isomers, the purified N-domain was cross-linked with DSS and the reaction products were analyzed by PAGE with or without SDS (Fig. 3). As can be seen in Fig. 3A, the cross-linked N-domain migrated in SDS-PAGE as five discrete bands possibly representing mono-, di-, tri-, tetra- and pentamer. Diffused bands migrating slower than the pentamer may represent artificial aggregation during the cross-linking reaction. The treatment of BSA with DSS under the same conditions did not cause any cross-linking (data not shown). In non-SDS-PAGE, cross-linked N-domain bands migrated faster than non-cross-linked ones (Fig. 3B), probably reflecting neutralization by the reaction with DSS of the positive charges on the side chain of Lys residues. The CEA N-domain carries four Lys residues (2). It should be noted that after cross-linking the number of visible bands increased to at least six which probably reflects some heterogeneity in the extent of the reaction.

Two-dimensional PAGE analysis. In order to see more clearly the relationship between the bands seen in the non-SDS PAGE and the degree of aggregation, two-dimensional PAGE of the cross-linked and non-cross-linked N-domain was performed. When non-cross-linked N-domain was electrophoresed without SDS in both directions, a diagonal pattern was observed (Fig. 4A). It suggested that no significant association or dissociation of the N-domain took place during the electrophoresis. On the other hand, when the electrophoresis in the second direction was performed in the presence of SDS, all the N-domain bands were found at the monomer position (Fig. 4B). A minor band seen in the middle of the gel probably represents contamination, more visible here than in one-dimensional SDS-PAGE that reflects enrichment during the first-dimension electrophoresis.

When the cross-linked N-domain was submitted to the two-dimensional PAGE under the same conditions as described in Fig. 4B, it was clearly seen that each band, from the faster to slower migrating ones in the first dimension consists, respectively, of mono-, di-, tri-, tetra- and pentamer (Fig. 4C). Apparent presence of multiple forms in each band, except for the fastest band, reflects the incompleteness of the cross-linking reaction.

DISCUSSION

The combined results of cross-linking and two-dimensional PAGE, clearly showed that recombinant N-domain of CEA, purified from *E. coli*, existed in multi-

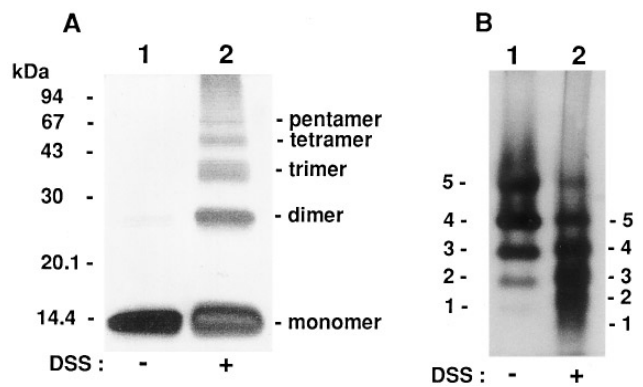

FIG. 3. Cross-linking of the recombinant CEA N-domain. Non-cross-linked (lanes 1) or cross-linked (lanes 2) CEA N-domains were analyzed by SDS-PAGE (15% gel) under reducing conditions (A) or PAGE (10% gel) without SDS (B). Cross-linking was performed with DSS as described in Materials and Methods.

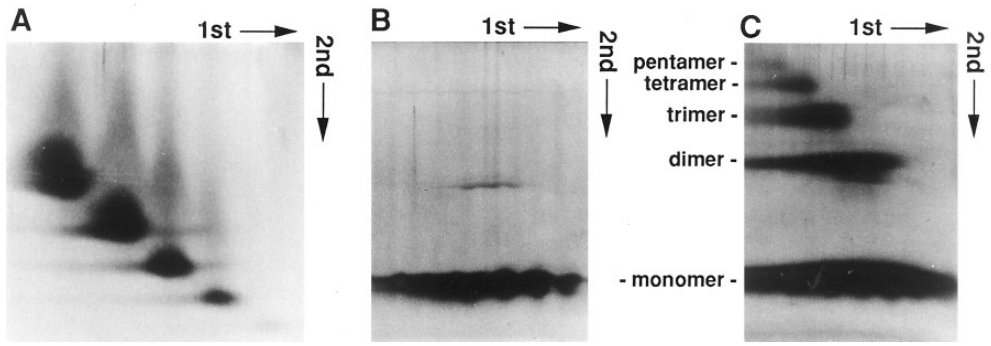

FIG. 4. Two-dimensional PAGE of non-cross-linked (A and B) and cross-linked (C) recombinant CEA N-domain. In A, both electrophoreses, in the first and the second direction, were performed without SDS. In B and C, electrophoreses in the first direction were performed without SDS followed by electrophoreses in the second direction with SDS. Cross-linking and PAGE conditions were the same as in Fig. 3.

ple oligomeric forms, ranging from monomer to pentamer. Apparently, the amount of monomer is the lowest, followed by dimer then trimer, the relative amounts of tri-, tetra- and pentamer being rather obscure because of the uncertainty in the relationship between the intensity of the staining and the number of molecules in these bands (Figs. 3B and 4A). The relative amounts of monomer and oligomers change drastically after cross-linking reaction probably because of the incompleteness of the reaction (Figs. 3A and 4C).

The oligomerization of N-domain observed in the present study by non-SDS PAGE for *E. coli* product is in good agreement with the results obtained by laser desorption/time of flight mass spectrometry for the purified N-domain expressed in HeLa cells (29). In HeLa cells N-domain was produced in a soluble, glycosylated form and could be regarded to assume conformation as near as possible to the natural one, except that it lacks the other domains which comprise major part of CEA. The N-domain was found in the form of monomer, dimer and trimer; the amounts being in the descending order, the existence of higher member multimers was not extensively pursued and thus obscure (29). The discrepancy in the relative amounts of monomer and multimers between *E. coli* and HeLa cell products may reflect the differences in experimental conditions such as solvents and concentrations of the proteins and/or in the degree of glycosylation as discussed below. These data clearly indicated that the self-association seemed to be of intrinsic feature of the N-domain rather than artifacts caused by conditions needed to solubilize and recover the recombinant products from *E. coli*. It was also suggested that sugar chains might not play important roles in the formation of multimers although a possibility remains that they rather interfere with formation of higher member multimers possibly by steric hindrance. The N-domain produced in *E. coli* is not glycosylated while the one produced in HeLa cells, like the natural one, seemed to be glycosylated at both of two potential N-glycosylation sites (29).

It is interesting to note that although the number of subunits in multimers of the N-domain increased up to five as evidenced by the clearly distinguishable bands in PAGE, the formation of hexamer and larger multimers was not clearly noticeable, if any (Figs. 3B and 4A). This is clear evidence that the multimers were not formed by non-specific aggregation which may produce also multimers consisting of more molecules. To explain this phenomenon, it is tempting to speculate the presence of two binding sites or regions "A" and "B" in the N-domain molecule: region "A" of one molecule

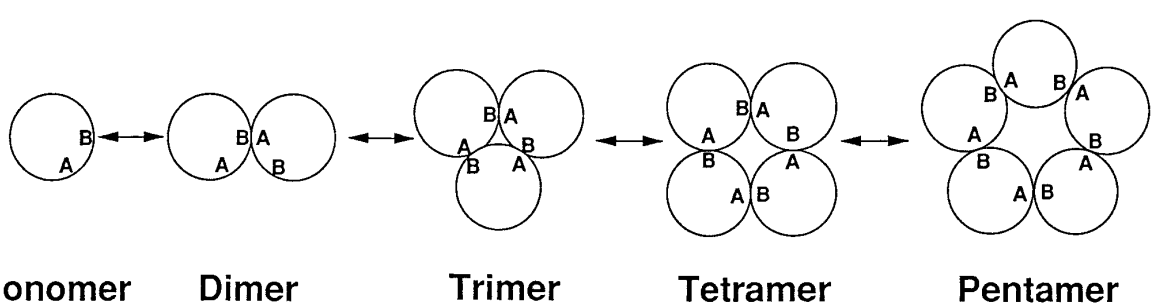

FIG. 5. Possible mode of self-assembling of CEA N-domain. The circles represent CEA N-domain molecules. "A" and "B" are possible binding sites that are responsible for oligomerization of the N-domain.

interacting with region "B" of another molecule as illustrated schematically in Fig. 5. The relative location of "A" and "B" might be so variable, possibly by virtue of flexibility of the N-domain structure, to allow the formation of oligomers up to pentamer but not hexamer or higher multimers.

It is tempting to speculate that the dimerization previously noted for CEA (14) is mediated by homophilic interaction between the N-domain, the oligomerization including more CEA strands being not favored because of the steric effects of other bulky domains, and the dimerization enhances homotypic, and/or heterotypic adhesion activities of CEA. It has been shown that ICAM exists in a dimer form and the soluble dimer form has two orders of magnitude higher affinity than the monomer form to its receptor, LFA-1 (30).

However, further detailed studies including the isolation and use of monomer form to clarify the kinetic features of the multimer formation and the experiments to see relationship of the N-domain-N-domain interaction and interaction between N-domain and 6 and/or 7th domain in homophilic adhesion of CEA as suggested (2531) are definitely needed.

ACKNOWLEDGMENT

This work was supported in part by Grant 6 PO4 A 01808 from the Committee of Scientific Studies (KBN Warsaw, Poland).

REFERENCES

1. Thompson, J. A., Grunert, F., and Zimmermann, W. (1991) *J. Clin. Lab. Anal.* **5,** 344–366.
2. Oikawa, S., Nakazato, H., and Kosaki, G. (1987) *Biochem. Biophys. Res. Commun.* **142,** 511–518.
3. Zimmermann, W., Ortlieb, B., Friedrich, R., and von Kleist, S. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 2960–2964.
4. Beauchemin, N., Benchimol, S., Cournoyer, D., Fuks, A., and Stanners, C. P. (1987) *Mol. Cell. Biol* **7,** 3221–3230.
5. Tawaragi, Y., Oikawa, S., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1988) *Biochem. Biophys. Res. Commun.* **150,** 89–96.
6. Watanabe, S., and Chou, J. Y. (1988) *Biochem. Biophys. Res. Commun.* **152,** 762–768.
7. Streydio, C., Swillens, S., Greorgs, M., Szpirer, C., and Vassart, G. (1990) *Genomics* **6,** 579–592.
8. Hinoda, Y., Neumaier, M., Hefta, S. A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L. J. F., Shively, J. E., and Raymond, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6959–6963.
9. Oikawa, S., Imajo, S., Noguchi, T., Kosaki, G., and Nakazato, H. (1987) *Biochem. Biophys. Res. Commun.* **144,** 634–642.
10. Paxton, R. J., Mooser, G., Pande, H., Lee, T. D., and Shively, J. E. (1987) *Proc. Natl. Acad Sci. USA* **84,** 920–924.
11. Takami, N., Misumi, Y., Kuroki, Mo., Matsuoka, Y., and Ikehara, Y. (1988) *J. Biol. Chem.* **263,** 12716–12720.
12. Hefta, S. A., Hefta, L. J., Lee, T. D., Paxton, R. J., and Shively, J. E. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4648–4652.
13. Williams, A. F., and Barcley, A. N. (1988) *Annu. Rev. Immunol.* **6,** 381–405.
14. Krop-Watorek, A., Sedlaczek, P., and Lisowska, E. (1989) *Arch. Immunol. Ther. Exp.* **37,** 703–714.
15. Krop-Watorek, A., Sedlaczek, P., and Lisowska, E. (1983) *Mol. Immunol.* **20,** 777–785.
16. Hunter, I., Sawa, H., Edlund, M., and Obrink, B. (1996) *Biochem. J.* **320,** 847–853.
17. Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K., and Stanners, C. P. (1989) *Cell* **57,** 327–334.
18. Oikawa, S., Inuzuka, C., Kuroki, Mo., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1989) *Biochem. Biophys. Res. Commun.* **164,** 39–45.
19. Oikawa, S., Kuroki, Mo., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1992) *Biochem. Biophys. Res. Commun.* **186,** 881–887.
20. Rojas, M., Fuks, A., and Stanners, C. P. (1990) *Cell Growth Differentiation* **1,** 527–533.
21. Neumaier, M., Paululat, S., Chan, A., Matthaes, P., and Wagener, C. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10744–10748.
22. Zhou, H., Fuks, A., and Stanners, C. P. (1989) *Cell Growth Differentiation* **1,** 209–215.
23. Eidelman, F. J., Fuks, A., DeMarte, L., Taheri, M., and Stanners, C. P. (1993) *J. Cell Biol.* **123,** 467–475.
24. Jessup, J. M., Kim, J. C., Thomas, P., Ishii, S., Ford, R., Shively, J. E., Durbin, H., Stanners, C. P., Fuks, A., Zhou, H., Hansen, H. J., Goldenberg, D. M., and Steele, G., Jr. (1993) *Int. J. Cancer* **55,** 262–268.
25. Oikawa, S., Inuzuka, C., Kuroki, Mo., Arakawa, F., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1991) *J. Biol. Chem.* **266,** 7995–8001.
26. Kuroki, M., Murakami, M., Wakisaka, M., Ikeda, S., Oikawa, S., Oshima, T., Nakazato, H., Kosaki, G., and Matsuoka, Y. (1992) *Immunol. Investigations* **21,** 241–257.
27. Laemmli, U. K. (1970) *Nature* **227,** 680–685.
28. Pilch, P. F., and Czech, M. P. (1979) *J. Biol. Chem.* **254,** 3375–3381.
29. Hefta, L. J. F., Chen, F. S., Ronk, M., Sauter, S. L., Sarin, V., Oikawa, S., Nakazato, H., Hefta, S., and Shively, J. E. (1992) *Cancer Res.* **52,** 5647–5655.
30. Reily, P. L., Woska, J. R., Jr., Jeanfavre, D. D., McNally, E., Rothlein, R., and Bormann, B-J. (1995) *J. Immunol.* **155,** 529–532.
31. Zhou, H., Fuks, A., Alcaraz, G., Bolling, T. J., and Stanners, C. P. (1993) *J. Cell Biol.* **122,** 951–960.