Nonhuman Cells Correctly Sort and Process the Human Lysosomal Enzyme Cathepsin D†

Gregory E. Conner,* Jenny A. Udey, Cristina Pinto,‡ and Juan Sola
Department of Anatomy and Cell Biology, University of Miami School of Medicine, Miami, Florida 33101
Received October 4, 1988; Revised Manuscript Received December 20, 1988

ABSTRACT: Cathepsin D, like most lysosomal enzymes, undergoes multiple proteolytic cleavages during its lifetime. Although the significance of the earliest cleavages of cathepsin D is apparent (loss of the NH_2-terminal signal peptide and activation peptide), functions of the two later cleavages are not understood and do not occur in all species. To examine these later events, a cDNA coding for human cathepsin D, which is normally processed to a two-chain form, was isolated and then expressed in mammalian cells from species which do not process the enzyme to the two-chain form. Analysis of the expressed human cathepsin D demonstrated proteolytic processing identical with that seen in normal human fibroblasts. Since processing to the two-chain form of the enzyme occurs in the lysosome, these studies revealed that the human cathepsin D was correctly sorted. The data also indicated that the sorting mechanism was conserved between diverse species and that late proteolytic processing in a variety of species was not determined by the presence or absence of the processing enzymes in the cell.

Cathepsin D (EC 3.4.23.5) is a major component of lysosomes in most higher eucaryotes. Due to its relative abundance, this enzyme has been purified from a variety of sources, characterized enzymatically [for a review, see Barrett (1977)], and sequenced by Edman degradation (Shewale & Tang, 1984; Takahashi & Tang, 1983; Erickson et al., 1981). A cDNA clone has also been isolated and sequenced (Faust et al., 1985). These studies have shown that cathepsin D is an aspartyl protease with an acid pH optimum. The enzyme bears striking amino acid sequence similarity and structural similarity with other aspartyl proteases which are extracellular enzymes (Tang, 1979) and thus provides a model system for studying the intracellular mechanism which distinguishes and targets lysosomal enzymes and secreted proteins to their correct destination.

Most lysosomal enzymes undergo multiple proteolytic cleavages during their life cycle in the cell. Cathepsin D has been well characterized in terms of the location of these cleavages within the primary structure (Conner et al., 1987). The enzyme is synthesized as a preproenzyme which is cleaved cotranslationally to yield the proenzyme. The proenzyme is presumed to be catalytically inactive until cleavage to the single- and two-chain forms observed in cells and tissue (Hasilik et al., 1982). These enzymatically active forms of cathepsin D are both metabolically stable, with the two-chain enzyme being derived from the single-chain enzyme by proteolysis (Huang et al., 1979). Comparison of the porcine cathepsin D protein sequence with the sequence deduced from the human kidney cDNA suggests that seven amino acids are removed at the cleavage site after generation of the two-chain enzyme. Finally at later times, approximately 1000 kDa are removed from the carboxy terminus of the molecule (Erickson & Blobel, 1983). These last two processing events have been shown to occur after the enzyme arrives in the lysosome (Hentze et al., 1984; Gieselmann et al., 1983).

The occurrence and relative amounts of single-chain and two-chain enzyme vary depending on the species. Rat, mouse, and hamster do not generate the two-chain form of the enzyme, human and pig ultimately process the majority of cathepsin D to the two-chain enzyme, and bovine tissue contains almost equivalent amounts of the two forms. The significance of this proteolytic cleavage is not known.

In order to study these later proteolytic events, we have cloned and sequenced a cDNA coding for human fibroblast cathepsin D. Using this cDNA molecule in transient expression experiments, we have demonstrated that rodent cells correctly sorted human fibroblast cathepsin D, but unlike their endogenous cathepsin D processed the human enzyme to the two-chain form. These studies showed that the proteolytic processing to the two-chain form of cathepsin D was not a function of the processing machinery but, instead, is most likely due to variations in the cathepsin D amino acid sequence among different species.

Experimental Procedures

Materials. [^{35}S]Methionine (800–1000 Ci/mmol) and [γ-^{32}P]ATP (5000 Ci/mmol) were purchased from Du Pont NEN Research Products. All tissue culture media and components were obtained from Grand Island Biological Co. Protein A–Sepharose was purchased from Pharmacia.

Isolation and Characterization of Cathepsin D cDNA. A human fibroblast cDNA library, sized to contain inserts 2000–2500 bp in length, was generously donated by Dr. H. Okayama (Okayama & Berg, 1983). The library, in *Escherichia coli* DH5, was screened by hybridization (Maniatis et al., 1982; Wood et al., 1985) using a synthetic oligonucleotide derived from the porcine spleen cathepsin D amino acid sequence (Takahashi & Tang, 1983). The oligonucleotide was synthesized on an Applied Biosystems DNA synthesizer (courtesy of Dr. Peter Model, The Rockefeller University) and was purified by chromatography on Bio-Gel P60. The oligonucleotide was labeled with [γ-^{32}P]ATP using T4 DNA kinase. Clones identified on the first screen were colony purified, rescreened with the oligonucleotide, and mapped with restriction endonucleases. Fragments of the longest clone, pCSPD1, were either transferred to M13 and sequenced using the dideoxy chain termination procedure (Sanger et al., 1977) or directly sequenced by the chemical method of Maxam and Gilbert (1980).

† This study was supported by NIH Grant GM35812 and by a grant from the Florida Division of the American Cancer Society (F84UM-3).
* To whom correspondence should be addressed.
‡ Present address: Instituto de Quimica Fisiologica, 1600 Lisboa, Portugal.

0006-2960/89/0428-3530$01.50/0

Cell Growth and Transfections. All cells were grown at 37 °C in 5% CO_2 in air. PtK2 cells (ATCC CCL56) and human skin fibroblasts were maintained in Ham's F10 and F12 medium, respectively. PK15 cells (ATCC CCL33) and mouse Ltk$^-$ cells were maintained in Dulbecco's modified Eagle's medium (DMEM).[1] Chinese hamster ovary (CHO) cells were maintained in α-modified Eagle's medium. All media were supplemented with 10% fetal bovine serum (FBS) and antibiotics.

Transfections were performed at room temperature in air. Cells, 50% confluent in 35-mm culture dishes, were washed 3 times in serum-free DMEM and then incubated for 1 h in serum-free DMEM containing 25 mM HEPES, pH 7.2, 200 μg/mL DEAE-dextran, and 4 μg/mL DNA (Sompayrac & Danna, 1981). Complete medium containing 10% dimethyl sulfoxide was added for 2–3 min, washed out, and replaced with fresh complete medium containing 8 or 80 μM chloroquine (PtK2 cells) for 1 or 4–5 h (PtK2 cells). Cells were washed and allowed to recover for 48 h before being labeled with [^{35}S]methionine.

Metabolic Labeling and Immunoprecipitation. For continuous labeling, cultures were placed in DMEM without methionine, but supplemented with 5% FBS and 200 μCi/mL [^{35}S]methionine for 12–16 h. Pulse labeling and chases were performed as described previously (Erickson et al., 1981). Cells were washed 3 times with phosphate-buffered saline and then harvested by solubilization in 0.5% SDS, 5 mM Tris, pH 8.2, 5 mM EDTA, 150 mM NaCl, and 0.1 mM phenylmethanesulfonyl fluoride. Extracts were heated to 100 °C for 5 min, sonicated, and centrifuged at 12000*g* for 10 min, and the supernatants were diluted with 10% Triton X-100 to a final concentration of 2.5%. Media samples were centrifuged at 12000*g* for 10 min and the supernatants transferred to a clean tube. Supernatants were made 0.5% in SDS, heated for 5 min at 100 °C, and then adjusted by the addition of Triton X-100 to 2.5% final concentration. An excess of anti-cathepsin D heavy-chain antibodies (Erickson et al., 1981) was added to the samples and allowed to react for 1 h at room temperature, followed by recentrifugation and transfer of supernatants to tubes containing protein A–Sepharose. Following a 1-h incubation on a rotator, the beads were washed 5 times in 2.5% Triton X-100, 0.5% SDS, 150 mM NaCl, 50 mM Tris, pH 8.2, 5 mM EDTA, and 0.25 M sucrose and 1 time in 150 mM NaCl, 10 mM Tris, pH 8.1, and 0.1 mM EDTA. Immunoprecipitates were eluted from the beads by heating for 5 min in 2% SDS, 10 mM Tris, pH 7.5, 5 mM EDTA, 20% sucrose, and 0.005% bromophenol blue. Electrophoresis through 12% acrylamide gels and fluorography have been previously described (Erickson et al., 1981).

Results and Discussion

Isolation of a Human Fibroblast cDNA Clone. A size-selected human fibroblast cDNA library was screened with a mixed-sequence oligonucleotide 29 bases in length. This oligonucleotide (Figure 1) was derived from amino acids 25–34 of porcine spleen cathepsin D (Takahashi & Tang, 1983), taking into account the codons used in the corresponding sequences of other aspartic proteases (Panthier et al., 1982; Imai et al., 1983; Sogawa et al., 1983; Moir et al., 1982). Two positive clones, identified after screening 50 000 colonies, were sized. The larger plasmid, pCPSD1, containing an insert 2101 bp in length, was mapped with restriction enzymes and sequenced. Comparison of the deduced amino acid sequence with the amino acid sequence of porcine cathepsin D (Shewale & Tang, 1984; Takahashi & Tang, 1983; Erickson et al., 1981) and with the sequence of human kidney cathepsin D (Faust et al., 1985) demonstrates that pCPSD1 contains the entire coding region for human fibroblast cathepsin D. The sequence of this cDNA differs from the human kidney cathepsin D by the 17-nucleotide-long sequence ACTGCGGCGTCATCCCG at the 5′ end of the cDNA and an A to G transition at base 1357. The extremely high level of nucleotide sequence similarity between kidney and fibroblast cDNAs isolated from different sources suggests a low level of allelic polymorphism at this locus in humans.

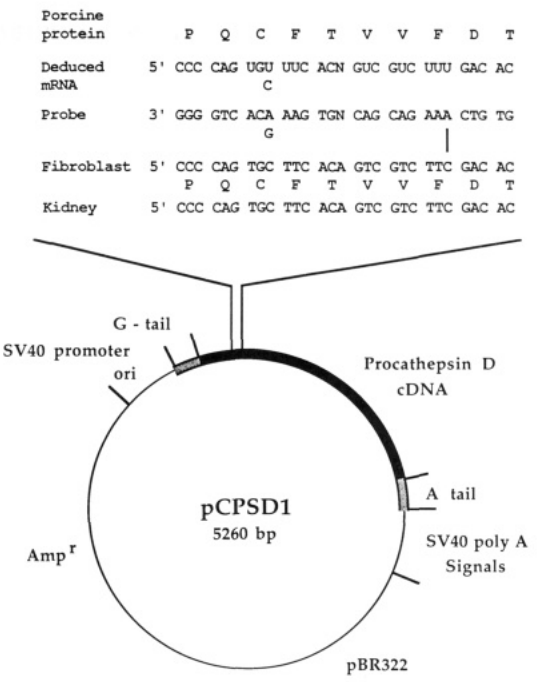

FIGURE 1: Isolation and identification of pCSPD1. A segment of amino acid sequence of porcine cathepsin D was used to derive a 29-residue oligonucleotide probe. Degeneracy was limited to eight possible sequences by following codon usage in other aspartic proteases (Panthier et al., 1982; Imai et al., 1983; Sogawa et al., 1983; Moir et al., 1982). The corresponding region of the human fibroblast sequence identified by hybridization with this probe demonstrates one mismatch in the second phenylalanine codon. This sequence is identical with that found in human kidney cathepsin D (Faust et al., 1985). The location of this sequence in pCPSD1 and salient features of the vector are also shown.

As isolated from the library, pCPSD1 is expressible in mammalian cells by virtue of the juxtaposed SV40 promoter. We have taken advantage of this property to explore species differences in postsynthetic processing of cathepsin D.

Expression of pCPSD1 in Nonhuman Cells. To study proteins expressed from transfected DNA molecules, it is necessary to distinguish the expressed protein from the endogenous host cell proteins. This is usually accomplished by transfecting cells which do not contain the protein of interest followed by identification of the product with antisera (e.g., viral proteins, differentiated products, etc.). Since many proteins with housekeeping functions are structurally conserved, they are frequently immunologically cross-reactive, thus making identification of the expressed protein with polyclonal antisera difficult. Alternative use of monoclonal antibodies complicates studies which involve making changes in the primary structure of the protein of interest. In order to study the expression of cathepsin D, a highly conserved and ubiquitous lysosomal protease, we have examined several cell lines

[1] Abbreviations: DMEM, Dulbecco's modified Eagle's medium; FBS, fetal bovine serum; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; CHO, Chinese hamster ovary.

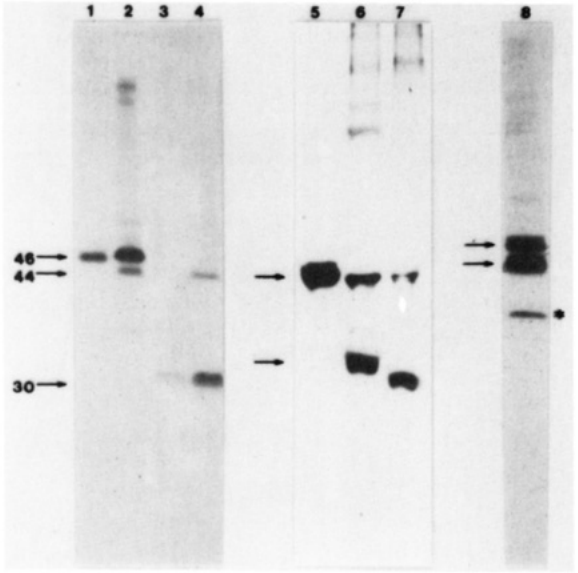

FIGURE 2: Comparison of endogenous cathepsin D from cultured cells. Cells were either pulse labeled and chased with cold methionine or continuously labeled with [^{35}S]methionine. Cultures were then solubilized, and the cell extracts and culture media were immunoprecipitated with anti-cathepsin D heavy-chain antibodies as described under Experimental Procedures. Immunoprecipitates were analyzed by electrophoresis and fluorography. Lanes 1 and 2, pulse-labeled PK15 and human fibroblast cathepsin D, respectively; lanes 3 and 4, pulse-labeled and chased PK15 and human fibroblast cathepsin D, respectively; lanes 5–8, continuously labeled Ltk$^-$, PK15, PtK2, and CHO cells, respectively. The asterisk denotes a major nonspecific protein found in the CHO immunoprecipitates. Arrows indicate the positions of the 46-kDa procathepsin D, the 44-kDa single-chain cathepsin D, and the 30-kDa heavy chain of cathepsin D.

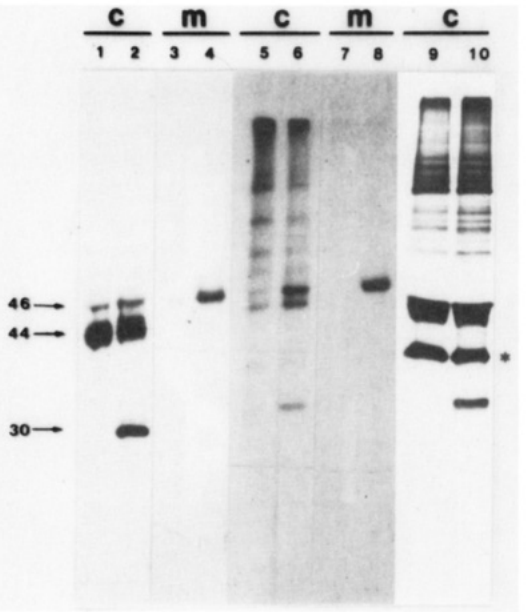

FIGURE 3: Immunoprecipitation of expressed human cathepsin D from nonhuman cells. After transfection with pCPSD1, and continuous labeling with [^{35}S]methionine, cells were solubilized, and the cell extracts and culture media were immunoprecipitated with anti-cathepsin D heavy-chain antibodies. Immunoprecipitates were analyzed by electrophoresis and fluorography. Lanes 1–4, mouse L cells; lanes 5–8, PtK2 cells; lanes 9–10, CHO cells. Lanes 1, 3, 5, 7, and 9 are from cells without transfection with pCPSD1; lanes 2, 4, 6, 8, and 10 are from cells after transfection with pCPSD1. C designates immunoprecipitates of cell extracts; M designates immunoprecipitates of culture media. Arrows indicate the positions of the 46-kDa procathepsin D, the 44-kDa single-chain cathepsin D, and the 30-kDa heavy chain of cathepsin D.

for differences in cathepsin D structure or metabolism which would allow their use as a host for transient expression assays.

Mouse Ltk$^-$, Chinese hamster ovary (CHO), and PtK2 (rat kangaroo) cell lines were chosen for transfection on the basis of differences in molecular mass between their endogenous cathepsin D and the human enzyme. Since the metabolism of porcine cathepsin D has been carefully analyzed in this laboratory, we compared the mobilities of cathepsin D from these cell lines with that of porcine PK15 and primary human fibroblast cultures (Figure 2). Cathepsin D was detected in these cell lines by immunoprecipitation of metabolically labeled cellular protein followed by gel electrophoresis and fluorography. Human and porcine cathepsin D had identical mobilities with both species processing cathepsin D to the two-chain enzyme represented by the 30-kDa heavy chain (Figure 2, lanes 1–4). The anti-cathepsin D antiserum was prepared against the 30-kDa heavy chain of porcine cathepsin D (Erickson et al., 1981) and clearly was cross-reactive with cathepsin D from a wide range of organisms (see below). The 15-kDa light chain was not seen because the antibody only recognizes the 30-kDa heavy chain, and immunoprecipitations were performed after SDS denaturation.

Ltk$^-$ and CHO cells only processed cathepsin D to the single-chain enzyme (Figure 2, lane 5 and lane 8). An additional band was seen in the CHO cell immunoprecipitate (asterisk, Figure 2, lane 8) which did not correlate with the anti-cathepsin D titer of the antiserum. PtK2 cells, like porcine and human cells, processed endogenous cathepsin D to the two-chain form (Figure 2, lane 7); however, the PtK2 heavy chain had a faster mobility on SDS gels when compared to the porcine (Figure 2, lanes 3 and 6) and human (Figure 2, lane 4) heavy chains.

After DEAE-dextran transfection and the subsequent transient expression of pCPSD1 in Ltk$^-$, CHO, and PtK2 cells, fully processed human cathepsin D was observed in all three cell lines and comigrated with authentic human fibroblast cathepsin D (Figure 3). In Ltk$^-$ and CHO cells, a band corresponding to the 30-kDa heavy chain of cathepsin D was seen only in cells which were transfected (Figure 3, lanes 2 and 10) and not in the control cells (Figure 3, lanes 1 and 9). Since [^{35}S]methionine labeling was continuous for 12 h before solubilizing the cells, immature forms of human cathepsin D were present, and thus the 44-kDa band probably contains both human and rodent single-chain enzyme. These data indicated that the 30-kDa protein seen after transfection was a result of expression and processing of human fibroblast cathepsin D.

After transfection and expression of pCPSD1 in PtK2 cells, a band of greater intensity and slightly slower mobility was observed in the 30-kDa region (Figure 3, lane 6) when compared to control cells (Figure 3, lane 5). This increased quantity and difference in mobility after transfection with pCPSD1 was consistent with the expression and processing of human cathepsin D heavy chain. Therefore, all three cell lines expressed and processed human fibroblast cathepsin D from the transfected pCPSD1. Assuming that the methionine contents of the human, mouse, and kangaroo cathepsin D are approximately equal (based on the sequence conservation between aspartyl proteases), these data indicated that human cathepsin D was synthesized at a rate equal to or greater than the endogenous enzyme.

Analysis of the immunoprecipitates obtained from the media conditioned during metabolic labeling (Figure 3, lanes 4 and 8) showed increased quantities of cathepsin D had been secreted after transfection. These data suggested that excess production of human cathepsin D resulted in saturation of the

cellular mechanism for sorting to the lysosome. Such a saturation of the sorting machinery should result in the secretion of other lysosomal enzymes although no measurement of this effect has been made in our system. The amount of cathepsin D expression varied significantly between experiments with the same cell line as did the ratio of intracellular human cathepsin D to the secreted forms (a mixture of human and rodent). Due to the variability of transient expression levels, stable transformants capable of expressing different levels of the enzyme will be necessary to study the effect of excess cathepsin D production on the ratio of intracellular to secreted forms as well as the effect on other lysosomal enzymes.

Processing of Human Cathepsin D in Rodent Cells. The cleavage of cathepsin D which results in formation of the two-chain enzyme is restricted to certain species. It has been shown to be sensitive to leupeptin, an inhibitor of cysteinyl proteases (Hentze et al., 1984), and appears to take place after the enzyme has been sorted to a dense lysosomal compartment (Gieselmann et al., 1983). The cleavage between the heavy and light chains of human cathepsin D occurs within a sequence which, by comparison with other aspartic proteases, is a nonhomologous insertion of approximately 11 amino acids. The experiments described here demonstrated that rodents, which do not normally cleave cathepsin D into two chains, clearly possessed the necessary machinery for cleavage of the human enzyme. This fact strongly implied that the third proteolysis of cathepsin D was dependent on some structural aspect of the enzyme which was different between the rodent and human proteins and that this difference was within the inserted sequence. These results contrast with recent studies of human kidney cathepsin D expression in *Xenopus* oocytes (Faust et al., 1987). In this system, processing of human cathepsin D to the two-chain form is not observed, presumably because oocytes lack a typical lysosomal compartment which would contain the necessary processing enzymes (Wall & Meleka, 1985).

The exact nature of the difference between mouse and human cathepsin D has not been determined. However, the primary structures of the rat, porcine, and bovine enzymes in the region of the inserted sequence have been recently determined (Yonezawa et al., 1988). These sequences together with the human sequence demonstrate that significant differences in the cleavage region do exist between species.

The three-dimensional structures of several secreted gastric proteases have been determined. Comparison of human cathepsin D with amino acid sequences of these proteases, and inspection of the structure of one of these enzymes, penicillopepsin (James & Sielecki, 1983), suggested that the sequence surrounding the cleavage site is a loop connecting two β strands and is located on the surface of the molecule. We suggest that this surface-disposed insertion contains the sequence difference between species which either do or do not perform the cleavage to the two-chain enzyme. We also propose that this sequence difference is responsible for determining the final form of the enzyme in the cell.

The inserted sequence of cathepsin D in which cleavage occurs is not found in the other secreted aspartic proteases and thus could potentially have a function in intracellular sorting of this aspartic protease to the lysosome. The sequence differences in this region of cathepsin D do not rule out this possibility. However, if this part of the molecule is involved in recognition by the sorting machinery, that machinery does not have a strict sequence requirement in a manner similar to the machinery which recognizes and translocates nascent chains across the endoplasmic reticulum.

ACKNOWLEDGMENTS

We thank Dr. Gunter Blobel for guidance and support during early parts of this work, Dr. Jordan Tang for sharing sequence data, Dr. Ann Erickson for the gift of antiserum to porcine heavy-chain cathepsin D, and Dr. Melanie Pratt for critical reading of the manuscript.

Registry No. Cathepsin D, 9025-26-7.

REFERENCES

Barrett, A. J. (1977) in *Proteinases in Mammalian Cells and Tissues* (Barrett, A. J., Ed.) pp 209–248, North-Holland Publishing Co., New York.

Conner, G. E., Blobel, G., & Erickson, A. H. (1987) in *Lysosomes: their role in protein breakdown* (Glaumann, H., & Ballard, J., Eds.) pp 151–162, Academic Press, London.

Erickson, A. H., & Blobel, G. (1983) *Biochemistry 22*, 5201–5205.

Erickson, A. H., Conner, G. E., & Blobel, G. (1981) *J. Biol. Chem. 256*, 11224–11231.

Faust, P. L., Kornfeld, S., & Chirgwin, J. M. (1985) *Proc. Natl. Acad. Aci. U.S.A. 82*, 4910–4914.

Faust, P. L., Wall, D. A., Perara, E., Lingappa, V. R., & Kornfeld, S. (1987) *J. Cell Biol. 105*, 1937–1945.

Gieselmann, V., Pohlmann, R., Hasilick, A., & von Figura, K. (1983) *J. Cell Biol. 97*, 1–5.

Hasilick, A., von Figura, K., Conzelmann, E., Nehrkorn, H., & Sandhoff, K. (1982) *Eur. J. Biochem. 125*, 317–321.

Hentze, M., Hasilick, A., & von Figura, K. (1984) *Arch. Biochem. Biophys. 230*, 375–382.

Huang, J. S., Huang, S. S., & Tang, J. (1979) *J. Biol. Chem. 254*, 11405–11417.

Imai, T., Miyazaki, H., Hirose, S., Hori, H., Hayashi, T., Kageyama, R., Ohkubo, H., Nakashishi, S., & Murakami, K. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7405–7409.

James, M., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Moir, D., Mao, J., Schumm, J. W., Vovis, G. F., Alford, B. L., & Taunton-Rigby, A. (1982) *Gene 19*, 127–138.

Okayama, H., & Berg, P. (1983) *Mol. Cell. Biol. 3*, 280–289.

Panthier, J.-J., Foote, S., Chambraud, B., Strosberg, A. D., Corvol, P., & Rougeon, F. (1982) *Nature 298*, 90–92.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Shewale, J. G., & Tang, J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3703–3707.

Sogawa, K., Fujii-Kuriyama, Y., Mizukami, Y., Ichihara, Y., & Takahashi, K. (1983) *J. Biol. Chem. 258*, 5306–5311.

Sompayrac, L. M., & Danna, K. J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7575–7578.

Takahashi, T., & Tang, J. (1983) *J. Biol. Chem. 258*, 6435–6443.

Tang, J. (1979) *Mol. Cell. Biochem. 26*, 93–109.

Wall, D. A., & Meleka, I. (1985) *J. Cell Biol. 101*, 1651–1664.

Wood, W. I., Gitschier, J., Laskey, L. A., & Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1585–1588.

Yonezawa, S., Takahashi, T., Wang, X. J., Wong, R. N. S., Hartsuck, J. A., & Tang, J. (1988) *J. Biol. Chem. 263*, 16504–16511.