Molecular Cloning and Primary Structure of Rat Testes Metalloendopeptidase EC 3.4.24.15†

Adrian Pierotti,‡,§ Ke-Wen Dong,‡ Marc J. Glucksman,‡ Marian Orlowski,‖ and James L. Roberts*,‡

The Dr. Arthur M. Fishberg Research Center for Neurobiology and Department of Pharmacology, Mount Sinai School of Medicine, One Gustave L. Levy Place, City University of New York, New York, New York 10029

Received June 4, 1990; Revised Manuscript Received August 8, 1990

ABSTRACT: The complete amino acid sequence of rat testes metalloendopeptidase (EC 3.4.24.15) was deduced from the nucleotide sequence of a cDNA clone isolated by screening a rat testes library with a polyclonal antibody raised against a homogeneous preparation of the rat testes enzyme. The correctness of the sequence was verified by N-terminal amino acid sequence analysis of the isolated enzyme and by partial amino acid sequence analysis of three tryptic peptides located near the N-terminus, the middle, and C-terminus of the native protein. The enzyme is composed of 645 amino acids with a molecular weight of 72985. This value is close to that of the purified rat testes and brain enzyme as determined by polyacrylamide gel electrophoresis under denaturing and reducing conditions and by molecular sieving chromatography. The enzyme contains the putative active-site sequence -H-E-F-G-H- that is homologous to the sequence in the active site of thermolysin and several other related bacterial enzymes, as well as to active-site sequences of several mammalian zinc metallopeptidases. No amino acid sequence homology, beyond this active site, was found with thermolysin, a bacterial zinc metalloendopeptidase, nor with several mammalian zinc metallopeptidases. Northern blot hybridization analyses showed the presence of mRNA encoding the enzyme in rat testes, but not in other rat tissues in spite of the finding that enzyme activity is widely distributed in all tissues and that relatively high activities are present in rat brain and pituitary.

Previous work in one of our laboratories led to the identification and isolation of a zinc metalloendopeptidase (EC 3.4.24.15) highly active in testes, brain, and pituitary (Orlowski et al., 1983, 1989). The enzyme (referred to here as endopeptidase 24.15, or EP 24.15) cleaves preferentially bonds on the carboxyl side of hydrophobic amino acid residues and differs in this respect from endopeptidase 24.11 ("enkephalinase"), another zinc metalloendopeptidase widely distributed in animal tissues, whose specificity is directed toward bonds on the amino side of hydrophobic residues (Kerr & Kenny, 1974). The finding that the enzyme does not cleave proteins and peptides having more than 20 amino acid residues, that it is highly active in peptide-producing tissues, and that it catalyzes the processing of some neuropeptides and degradation of others indicated a function in the metabolism of bioactive peptides. Several enkephalin-containing peptides such as α- and β-neoendorphin, dynorphin$^{1-8}$, methionine-enkephalin-Arg-Gly-Leu, and metorphamide were rapidly converted by the enzyme to the corresponding enkephalins (Chu & Orlowski, 1985; Acker et al., 1987; Orlowski et al., 1989). The enzyme also cleaves the central Tyr^5–Gly^6 bond in gonadotropin releasing hormone (GnRH) (Orlowski et al., 1983) and constitutes the primary factor responsible for degradation of GnRH in membrane preparations from the hypothalamus and pituitary (Molineaux et al., 1988). Recent work has shown that in vivo inhibition of the enzyme by specific, active-site-directed inhibitors (Chu & Orlowski, 1984; Orlowski et al., 1988) increases the half-life of administered GnRH both in brain and in peripheral tissues (Lasdun et al., 1989) and greatly increases the magnitude and duration of secretion by the pituitary of luteinizing hormone (LH) and follicle stimulating hormone (FSH) (Lasdun & Orlowski, 1990).

Endopeptidase 24.15 exists in two forms in tissues and cell lines: a predominant soluble form constituting about 80% of the total activity and a minor membrane-bound form accounting for the remainder of activity (Acker et al., 1987). The two forms of EP 24.15 are very similar if not identical with respect to substrate specificity, immunological properties, and sensitivity to specific inhibitors. It is not yet known, however, whether the two forms differ by virtue of the presence in the membrane-bound form of a membrane-spanning segment, as seen in several zinc-containing peptidases, or by virtue of posttranslational modifications segregating the enzyme into different compartments of the cell.

As part of an effort to characterize further the two forms of EP 24.15, we used specific polyclonal antibodies raised in rabbits against the soluble form of the enzyme to screen a rat testes cDNA library and have as a result identified several cDNA clones encoding the enzyme. Here we present the primary structure of the enzyme predicted from the nucleotide sequence of the cDNA representing the mRNA encoding EP 24.15. We also report on the tissue distribution of this mRNA and the results of some experiments designed to compare the native enzyme from rat testes with that expressed as a recombinant protein in bacteria.

Materials and Methods

Isolation of the Enzyme and Protein Sequencing. A homogeneous preparation of EP 24.15 isolated from rat testes by a procedure previously described was used for raising antibodies in rabbits (Orlowski et al., 1989). For peptide se-

† This work was supported by Grants DK 25377 (M.O.) and DK 39029 (J.L.R.) from the National Institutes of Health.

* Address correspondence to this author at The Arthur M. Fishberg Research Center for Neurobiology, Mount Sinai School of Medicine, Fifth Ave. and 100th St., New York, NY 10029.

‡ The Dr. Arthur M. Fishberg Research Center for Neurobiology.

§ Present address: University Pierre Et Marie Currie, Group de Neurobiochimie, Cellulaire et Moleculaire, 96 Boulevard Raspail, Paris, France 75006.

‖ Department of Pharmacology.

0006-2960/90/0429-10323$02.50/0

quencing, enzyme solutions containing about 100 μg of protein/mL were concentrated in Centricon 10 microconcentrators (Amicon, Danvers, MA) to about 100 μL, and the protein was precipitated by addition of 100% trichloroacetic acid to a final concentration of 10% and placed on ice for about 30 min. The suspension was centrifuged, and the precipitated protein was washed twice with cold acetone and dried. The protein was then dissolved in 50 μL of 8 M urea containing 0.4 M ammonium bicarbonate, and after addition of 5 μL of 45 mM dithiothreitol and incubation at 50 °C for 15 min it was treated with 5 μL of 100 mM iodoacetamide. Tryptic digestion was then carried out after addition of water (140 μL) and trypsin in a ratio of 1:25 of enzyme to protein. HPLC separation of tryptic peptides was carried out as described by Stone et al. (1989). Amino acid sequence analysis of tryptic peptides was carried out on an Applied Biosystems 470 gas-phase protein sequencer. Three tryptic peptides were sequenced and are referred to as peptides 1, 2, and 3 in increasing order of their distance from the N-terminus of the protein. The repetitive yield of Val residues (5, 11) in peptide 1 was 90.3%, and the respective yield of Asp residues (2, 10) in peptide 3 was (89.8%).

Determination of Enzyme Activity. Enzyme activity was determined with Bz-Gly-Ala-Ala-Phe-pAB[1] as the substrate (0.8 mM) in a coupled enzyme assay in the presence of a large excess of aminopeptidase N. The enzyme cleaves the Gly–Ala bond of the substrate, and the aminopeptidase present in the incubation mixture cleaves the product of the reaction (Ala-Ala-Phe-pAB) to the constituent amino acids and free pAB. The aromatic amine is then determined after diazotization. Studies on substrate specificity showed a requirement for a minimum of five peptide bonds in synthetic substrates of the type used here for activity determination. While replacement of Gly in Bz-Gly-Ala-Ala-Phe-pAB with an aromatic amino acid residue both increases the rate of reaction and lowers the K_m, this substrate is nevertheless used for routine activity determinations because of its good solubility, and the relative simplicity of its synthesis (Orlowski, 1983). The conditions for activity measurements were the same as those described previously (Orlowski et al., 1983), except that determinations were carried out in a single-step reaction with the aminopeptidase present in the initial incubation mixture. A unit of enzyme activity is defined as the amount of enzyme that releases 1 μmol of the substrate/h. Specific activity is expressed in term of units per milligram of protein as determined by the method of Lowry et al. (1951).

Determination of Inhibition Constants. Substrate-related inhibitors of EP 24.15 (Orlowski et al., 1988) were used to determine their effect on the activity of the enzyme expressed in bacteria. K_i determinations were carried out by the method of Dixon (1959) at three different substrate concentrations and six different inhibitor concentrations (plots of $1/v$ versus inhibitor concentration) using a computer program. Coefficients of determination (r^2) were generally better than 0.99.

Isolation and Characterization of cDNA Clones Encoding EP 24.15. A rat testes cDNA library utilizing the λZapII vector containing 1.5×10^6 recombinants was purchased from Stratagene (La Jolla, CA). The library was screened as follows: 200 μL of an overnight culture of *Escherichia coli* XL-1 blue cells was mixed with an aliquot ($\sim 5 \times 10^4$ pfu/plate) of the library. After preabsorbtion for 15 min at 37 °C, the cells were plated onto NZ amine media plates. The plates were incubated at 42 °C for 3 h, transferred to 37 °C for an additional hour, and then overlayed with a nitrocellulose filter previously saturated with a 10 mM solution of IPTG. The plates with the filter were incubated for an additional 3 h at 37 °C. The filters were then removed and washed briefly in TBST (10 mM Tris, pH 8.0, 150 mM NaCl, and 0.05% Tween 20) and stored in a fresh solution of TBST overnight.

An antibody raised against a homogeneous preparation of EP 24.15 (Orlowski et al., 1989) was used as a probe. All further manipulations were carried out at room temperature. The filters were incubated in TBST containing 10% horse serum for 30 min to block nonspecific binding and then incubated for 1 h with a 1/2000 dilution of the primary antibody in TBST containing 0.5% BSA. The filters were washed 3 times in TBST for 10 min and transferred to a 1/3000 dilution of goat anti-rabbit IgG–alkaline phosphatase conjugate in TBST (Bio-Rad, Richmond, CA). The filters were incubated for 1 h and washed 3 times in TBST for 10 min. To visualize positive colonies, filters were transferred to the color development solution consisting of 330 μg/mL NBT and 165 μg/mL bovine calf intestine alkaline phosphatase substrate in alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.5, 100 mM NaCl, and 5 mM $MgCl_2$). The reaction was allowed to continue for 45 min before the filters were washed in distilled water to stop the reaction. Positive phage plaques were isolated and rescreened to obtain single positive phage isolates. The phage DNA was prepared according to the manufacturer's recommendation (Stratagene, San Diego, CA). The inserted DNA fragments were subcloned into phage M13-mp 18 and -mp 19. The single-stranded DNA serial deletion procedure, described by Dale et al. (1985), was used to produce a sequential series of overlapping clones for use in DNA sequencing. DNA was sequenced by the chain terminator dideoxy method of Sanger (1977) using T7 polymerase. Because in some regions of the DNA sequencing in both directions could not give an unambiguous sequence, a set of synthetic 20-base oligonucleotides was synthesized and used for dideoxy sequencing (Figure 1, dotted arrows).

Expression of EP 24.15 in Bacteria. Two hundred microliters of overnight bacterial cultures containing the plasmid pBluescript with or without a cloned insert was grown in 5 mL of medium with antibiotic selection to an absorbance reading of 0.6 at 600 nm. IPTG was then added to a final concentration of 6 mM, and the bacteria were grown for an additional 3 h. The bacterial suspension was centrifuged at room temperature for 5 min at 3000*g*, and the pellet was resuspended in 250 μL of 50 mM Tris-HCl, pH 6.7. Bacteria were then lysed by three cycles of freezing and thawing followed by sonication. Bacterial debris was removed by centrifugation (15000*g* for 10 min), and the supernatant was assayed for enzyme activity as described. The non-insert-containing plasmid system was used as a control.

Western and Northern Blotting. Ten-microgram aliquots of protein from bacterial extracts were electrophoresed under denaturing conditions in 8% acrylamide gels as described (Laemmli, 1970) using a Protein II apparatus (Bio-Rad, Richmond CA). Electrophoretic transfer of proteins from SDS–polyacrylamide gels onto nitrocellulose was carried out essentially as described (Towbin et al., 1979; Burnette, 1981) in a Trans-Blot electrophoretic transfer cell (Bio-Rad, Richmond, CA). Visualization of proteins using anti-EP 24.15 antibodies was carried out as described for bacterial filters. Ten micrograms of RNA isolated by the guanidinium thiocyanate/lithium chloride procedure (Maniatis et al., 1982) from rat testes, lung, spleen, brain, cardiac muscle, adrenals,

[1] Abbreviations: BSA, bovine serum albumin; Bz, benzoyl; IPTG, isopropyl thiogalactoside; NBT, nitro blue tetrazolium; pAB, *p*-aminobenzoate.

Table I: Specific Activity of the Cloned Enzyme As Compared with Activity in Rat Tissues[a]

source of enzyme	activity [mmol (mg of protein)$^{-1}$ h^{-1}]
bacterial enzyme	56.8 (7100)
testis	4.03 (504)
brain	0.80 (100)
spinal cord	0.65 (81)
anterior pituitary	0.59 (74)
spleen	0.59 (74)
lung	0.31 (39)
kidney cortex	0.26 (33)
liver	0.25 (31)
skeletal muscle	0.24 (30)
heart	0.18 (23)
adrenal	0.11 (14)

[a] Activity was determined in supernatants of tissue homogenates prepared from mature male Sprague Dawley rats and in the soluble fraction of bacterial extracts using Bz-Gly-Ala-Ala-Phe-pAB as described under Materials and Methods. Data are mean values obtained from two separate determinations. Values in parentheses represent relative activity with respect to the activity in rat brain arbitrarily set at 100.

kidney, and liver was used for Northern blot analysis. RNA electrophoresis (Maniatis et al., 1982) was carried out on denaturing (1% formaldehyde) 0.8% agarose gels and was followed by transfer to BA85 nitrocellulose (Schleicher & Schuell, Keen, NH). The nitrocellulose filters were prehybridized for 3 h, hybridized with random-primed EP 24.15 cDNA probe for 18 h, washed, and UV-cross-linked under standard conditions. The positions of 18S and 28S rRNAs were determined by ethidium bromide staining of the gel before blotting. This revealed no gross RNA degradation in the different RNA samples. To verify that the mRNA in the samples was intact, the blot was stripped and reprobed with an RNA probe to cyclophillin mRNA, a ubiquitous mRNA. Intact cyclophillin mRNA was seen in all samples (data not shown), implying that the mRNA was intact.

Results

Screening of 3×10^6 recombinants, approximately twice the library, identified 28 immunopositive plaques, which survived subsequent rounds of screening and purification. The intensity of immune-positive reactions varied in different clones expressing EP 24.15. Several positive clones representing weakly and strongly positive immunoreactive plaques were grown in separate cultures and used for preparation of bacterial extracts which were then examined for EP 24.15 activity and immunoreactivity by Western blotting after separation by polyacrylamide gel electrophoresis. In general, the stronger immunoreactive plaques produced bacteria that yielded extracts giving more intense immunoreactive bands in Western blots, and a corresponding higher activity in enzymatic assays. It should be noted that the electrophoretic mobility of the immunoreactive proteins derived from different clones was very similar to or identical with that of the native enzyme isolated from rat testes, suggesting that the two enzymes had very similar molecular weights.

One of the clones (C57) expressing high enzyme activity was grown in a 500-mL culture, and after isolation of the bacteria, and preparation of a soluble extract, EP 24.15 activity was determined as described under Materials and Methods. The specific activity of the extract was compared with that determined in the soluble fraction of homogenates prepared from different rat tissues. The results shown in Table I indicate that unusually high amounts of active EP 24.15 are being expressed in bacteria. Thus, the specific activity of bacterial extracts was about 14 times higher than in rat testes, a tissue with the highest EP 24.15 activity, and about 70 times higher than in brain, a tissues with the second highest activity among the tissues listed in Table I. It should be noted that extracts from control bacteria grown under similar conditions had no EP 24.15 activity. Assuming that the specific activity of the bacterial enzyme is the same as that of the isolated testes enzyme (1220 μmol mg^{-1} h^{-1}; Orlowski et al., 1989), it can be calculated that an approximately 21–22-fold purification would be sufficient for the isolation of a homogeneous enzyme from bacterial extracts and that the enzyme may constitute up to 4.5% of the total protein of the soluble fraction of bacterial extracts. This suggests that bacterial clones might become a convenient source of large amounts of EP 24.15.

Table II: Effect of Active-Site-Directed Inhibitors of Endopeptidase 24.15 on the Cloned Bacterial Enzyme[a]

inhibitor	K_i (μM)	
	bacterial enzyme	testes enzyme
cFP-AAF-pAB	0.021 ± 0.002	0.027[b] ± 0.003
cFP-AAY-pAB	0.019 ± 0.002	0.016[b] ± 0.007

[a] Data are mean values ± SEM obtained from 10 separate determinations. [b] Values taken from Orlowski et al. (1988).

To characterize the expressed enzyme, the inhibitory constants of two specific active-site-directed inhibitors of EP 24.15 (Orlowski et al., 1988) were tested toward the bacterial enzyme. The two inhibitors, *N*-[1(*RS*)-carboxy-3-phenylpropyl]-Ala-Ala-Phe-*p*-aminobenzoate (cFP-AAF-pAB) and *N*-[1(*RS*)-carboxy-3-phenylpropyl]-Ala-Ala-Tyr-*p*-aminobenzoate (cFP-AAY-pAB), are substrate-related specific inhibitors of endopeptidase 24.15. As shown in Table II, the K_i values for both enzymes were virtually identical. It should be noted that *N*-[1(*RS*)-carboxy-3-phenylpropyl]-Phe-*p*-aminobenzoate (Pozsgay et al., 1986), an inhibitor of endopeptidase 24.11 ("enkephalinase"), another zinc metalloendopeptidase widely distributed in rat tissues, had no effect on the activity of the bacterial enzyme at concentrations exceeding 3 orders of magnitude of those used in experiments with the EP 24.15 inhibitors.

Amino Acid and Nucleotide Sequences of EP 24.15. One of the largest cDNAs (C57) containing 2.4 kb was sequenced in both directions (the sequence strategy is shown in Figure 1) and found to contain a large open reading frame. Because of the absence of the major consensus polyadenylation site, a poly(A) tail in the 3′ region of the cDNA, the absence of a clear initiator methionine, and a hydrophobic leader sequence as would be expected for a secreted or membrane-bound protein, other clones were examined for the presence of additional sequences at both the 5′ and 3′ ends. By use of a restriction map of clone C57, additional clones were analyzed by digestion with the restriction enzymes *Kpn*I and *Stu*I. This resulted in the identification of an additional clone (C44) that contained extensions in both the 3′ and 5′ direction. This clone was used for sequencing the 3′ and 5′ extensions and also for internal primed sequencing aimed at resolving ambiguities present in the sequence obtained from clone C57. As a result, the cDNA sequence shown in Figure 2 was obtained. To verify the position of the 5′ end of the EP 24.15 mRNA, a synthetic oligomer complementary to the 5′ end of the deduced sequence obtained from clone C57 was used in a primer extension analysis of testes mRNA. This analysis suggested the presence of RNA extensions 28, 75, and 135 nucleotides from the 5′ end of the presented sequence (data not shown), suggesting that we had a close to full-length cDNA. The presence of poly(A) sequence at the 3′ end of clone C44 clearly shows that

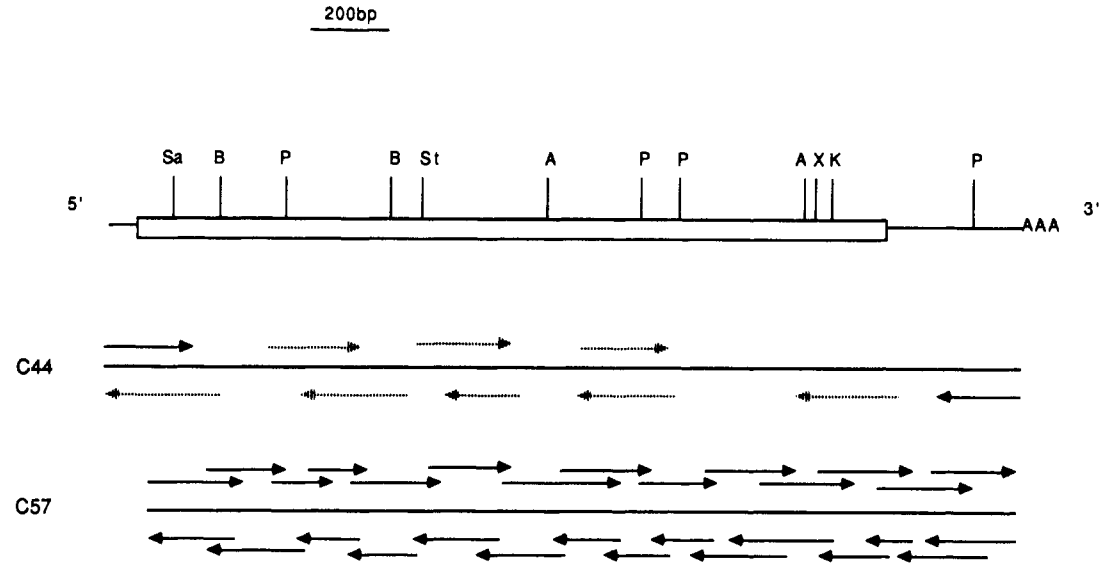

FIGURE 1: Sequencing strategies for clones C44 an C57 are shown. Solid arrows represent regions sequenced with universal M13 primers homologous to the vector, and the dotted arrows represent regions sequenced by using internal oligomers homologous to EP 24.15 cDNA. The top part of the figure represents the composite mRNA structure showing the coding region. Selected restriction endonuclease sites in the cDNA are indicated (A = *Acc*I, B = *Bst*EII, K = *Kpn*I, P = *Pst*I, Sa = *Sac*I, St = *Stu*I, and X = *Xho*I).

the infrequently used polyadenylation recognition sequence AUUAAA is used in EP 24.15 mRNA.

The longest open reading frame of the cDNA encodes a protein with 645 amino acid residues (Figure 2). The molecular weight of the protein calculated from the amino acid composition is 72 985, a value very close to that found for the isolated rat testes enzyme (70 000) as determined by polyacrylamide gel electrophoresis under denaturing and dissociating conditions (Orlowski et al., 1989), and also by molecular sieving chromatography for the brain enzyme (Orlowski et al., 1983). The protein has a predominance of acidic over basic amino acid residues (85 versus 79). Initial N-terminal amino acid sequence analysis by automated Edman degradation of the isolated rat testes enzyme presented difficulties because of the small amounts of available homogeneous enzyme, poor yield in the initial cycles, and difficulties in conclusive amino acid assignment in some cycles. The following sequence was, however, obtained: X_1-X_2-Pro-Ala-Ala-X_3-X_4-Gly-Asp-Val-Val with the X (dotted line) designating cycles in which it was not possible to make an unambiguous amino acid assignment, as more than one amino acid was present. Nonetheless, in cycles X_1 and X_2, the presence of Lys and Pro, respectively, was confirmed; similarly, cycles X_3 and X_4 contained Cys and Ala. On the basis of these findings, it can be assumed that the ATG codon immediately upstream of this sequence (Figure 2) codes for the first amino acid in the protein and that apparently the initiating Met is subsequently enzymatically cleaved, leaving the Lys residue as the amino terminus. Three peptides isolated by HPLC from tryptic digests of the native enzyme were subjected to automated Edman degradation. Their positions in the protein are underlined. For several residues (marked with a dotted line), identification was not unequivocal, because of the presence of some secondary sequences that could not be removed by HPLC repurification of the separated peptides. The position, however, of the three peptides, one near the N-terminus, the second in the middle of the protein, and the third near the C-terminus, confirms that the open reading frame of the cDNA encodes EP 24.15.

EP 24.15 had no signal sequence, and at present the manner in which the membrane-bound from the enzyme in anchored is not known. The molecule is generally hydrophilic with a

```
                                                  M K P P A A      6
GGACCGCACAGGCGTGAGGACAGAGGAGGGCGTGCCGTCCCGCGCGCAGCTGCCGCCATGAAACCCCCCGCAGCC   75
C A G D V V D T V S P C S T V N H L R W D L S A Q                 31
TGTGCTGGGGACGTGGTGGATACAGTGTCACCATGCTCCACCGTGAATCATCTGCGCTGGGACCTGAGCGCACAG  150
Q I R A L T T Q L I E Q T K C V Y D R V G A Q D F                 56
CAGATCAGAGCTCTAACCACACAGCTCATCGAGCAGACCAAGTGTGTGTACGACCGTGTGGGCGCCCAGGACTTC  225
E D V S Y E S T L K A L A D V E V T Y T V Q R N I                 81
GAGGACGTGTCCTATGAGAGCACACTGAAGGCACTGGCTGACGTGGAGGTCACCTACACAGTGCAGAGGAACATT  300
L D F P Q H V S P N K D I R A A S T E A D K K L S                106
CTCGACTTCCCCCAGCACGTGTCTCCGAACAAGGACATCCGCGCAGCCAGCACAGAGGCTGACAAGAAGCTCTCA  375
E F D V E M S M R Q D V Y Q R V V W L Q E K I P K                131
GAGTTTGATGTGGAGATGAGCATGAGGCAGGACGTGTACCAGAGGGTCGTGTGGCTGCAGGAGAAAATCCCGAAA  450
D S L K P E A A R Y L E R L I K L G R R N G L H L                156
GACTCACTGAAGCCTGAGGCAGCTCGCTACCTGGAGCGGCTCATCAAGCTGGGCCGGAGAAACGGGCTCCACTTA  525
P Q D T Q E K I K N I K K R L S L L C I D F N K N                181
CCTCAGGACACACAGGAGAAGATCAAGAACATCAAGAAGAGGCTGAGCCTGCTGTGCATCGACTTCAACAAGAAC  600
L N E D T T F L P F T R E E L G G L P E D F L N S                206
CTGAATGAGGACACCACCTTCCTGCCCTTCACGAGAGAGGAGCTGGGCGGGCTCCCCGAGGACTTCCTGAACTCC  675
L E K T E D G K L K V T L K Y P H Y F P L L K K C                231
CTGGAGAAGACAGAGGACGGCAAACTGAAGGTCACCCTCAAGTATCCACACTATTTCCCACTGCTGAAGAAGTGC  750
H V P E T R R L L E E A F N C R C K E E N C A I L                256
CACGTGCCCGAGACACGACGCCTGTTGGAGGAGGCCTTCAACTGTCGCTGCAAGGAGGAGAACTGTGCCATCCTG  825
K E L V S L R A Q K S N L L G F R T H A D Y V L E                281
AAGGAGCTAGTGTCCCTGCGGGCGCAGAAGTCCAACCTGCTGGGATTCCGCACACACGCAGACTACGTCCTGGAG  900
M N M A K T S Q T V A T F L D E L A R K L K P L G                306
ATGAACATGGCCAAGACCAGTCAGACAGTAGCCACCTTCCTAGATGAACTGGCCCGGAAGCTGAAGCCGCTGGGT  975
E Q E R A V I L E L K E A E C Q A R A A I R W A H                331
GAGCAGGAGCGTGCGGTGATCCTGGAGCTGAAGGAGGCAGAGTGCCAAGCGCGGGCTGCCATTCGATGGGCGCAT 1050
P R M G H A V L H E P G G G G T R Y R V D Q N L L                356
CCACGCATGGGACATGCGGTACTACATGAACCAGGTGGAGGAGGAACACGCTACCGCGTAGACCAGAACCTGCTG 1125
K E Y F P M Q V V T R G L L A I Y Q E L L G L T F                381
AAGGAGTACTTCCCCATGCAGGTGGTCACGCGTGGGCTGCTGGCCATCTACCAAGAGCTGCTGGGGCTGACCTTC 1200
T L E E G A A A W H E D V R L Y S V R D A A S G E                406
ACGCTGGAGGAAGGTGCTGCCGCCTGGCACGAAGACGTGCGGCTGTACTCTGTGCGTGACGCCGCCTCTGGAGAG 1275
E I G K F Y L D L Y P R E G K Y G H A A C F G L Q                431
GAGATTGGCAAGTTCTACCTCGACCTGTACCCCAGGGAAGGGAAGTATGGCCACGCGGCCTGCTTCGGCCTGCAG 1350
P G C L R Q D G S R Q L A I A A M V A N F I K P T                456
CCTGGCTGCCTACGGCAGGATGGCAGCCGACAGCTGGCCATCGCGGCCATGGTGGCCAATTTCACCAAGCCCACA 1425
P D V P S L L Q H D E V E T Y F H E F G H V M H Q                481
CCTGACGTGCCTTCCCTGCTGCAGCACGATGAGGTGGAGACCTACTTCCACGAGTTCGGGCACGTCATGCACCAG 1500
L C S Q A E F A M F S G T H V E R D F V E A P S Q                506
CTCTGCTCACAGGCAGAGTTTGCTATGTTCAGTGGGACCCACGTGGAGCGGGACTTTGTGGAGGCACCGTCACAG 1575
M L E N W V W E K E P L M R M S Q H Y R T G G E A                531
ATGCTGGAGAACTGGGTGTGGGAAAAGGAGCCACTGATGCGCATGTCCCAGCATTACCGCACAGGAGGCGAGGCC 1650
P E D L L E K L I K S R Q A N A G L F N L R Q I V                556
CCTGAGGACCTCCTGGAGAAGCTCATCAAGTCTCGCCAGGCCAATGCAGGTCTCTTCAACCTTCGGCAGATTGTC 1725
L A K V D Q V L H T Q T D V D P A E E Y A R S A R                581
CTTGCCAAGGTAGATCAGGTCCTGCACACACAGACAGATGTAGACCCAGCTGAGGAATATGCCCGCTCTGCCAGG 1800
R S L G C Q P H Q V P T C Q P L L A T S L V A T T                606
AGATCCTTGGGGTGCCAGCCACACCAGGTACCAACATGCCAGCCACTTTTGGCCACCTCGCTGGTGGCTACGACG 1875
L S T M A T C G V R C T P W T C S T H A S S R R V                631
CTCAGTACTATGGCTACTTGTGGAGTGAGGTGTACTCCATGGACATGTTCCACACACGCTTCAAGCAGGAGGGTG 1950
C L A P R L A W I T G P A S                    645
TGCTTAGCCCCAAGGTTGGCATGGATTACCGGACCAGCATCCTGAGGCCGGGGGGCTCTGAGGATGCCAGTACCA 2025

TGCTGAAGCAGTTCCTGGGCCGTGACCCCAAGCAAGACGCCTTCCTCCTGAGCAAGGGTCTGCAGGTAGAGGGCT 2100

GCGAGCCACCTGCGTGCTGAGCGGTCCCCAGCTCGCCCTCCTGTGCCTTAGCCTGCCGGGCGGCAAGGTCAGCAG 2175

CCCTTGGGTACGGATGAGCCTGGGGCAGTTTCCATCCTCGTCCTCCCCCCAGCTCCAGGTCCCCGGCTTCAGGCT 2250

GCTGTGCTCTCCTTGCACTATCCCCTCCCCAAACTGATGTTGGACATTAAATTTCACACAGGGGAAAAAAAAAAA 2325

AAAAAAA
```

FIGURE 2: Nucleotide sequence of the cDNA encoding endopeptidase 24.15 and the amino acid sequence of the enzyme obtained from the open reading frame. Amino acids are numbered in boldface beginning with N-terminus. The sequences of the N-terminal peptide and three tryptic peptides obtained by automated Edman degradation are underlined. Residues where amino acid assignment was not unambiguous are underlined with a dotted line. A single consensus glycosylation site is underlined with a broken line. The catalytic site containing two Zn^{2+}-coordinating histidine residues and a glutamate residue is double-underlined. The polyadenylation sequence near the 3′ end is also underlined. For further details, see the text.

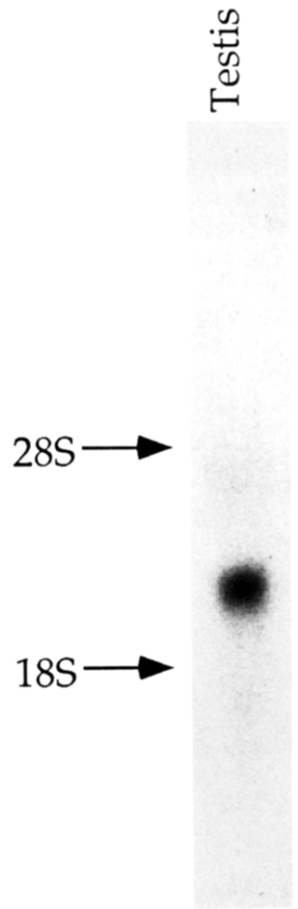

FIGURE 3: Total cellular RNA (10 μg) from the testis was applied for Northern blot analysis. Positions of the 18S and 28S rRNAs are indicated by the solid arrows.

short stretch of hydrophobic residues present around position 450 and a less hydrophobic domain around position 600. A single consensus sequence, Asn-Phe-Thr (Figure 2), that could represent an N-glycosylation site is present at position 451; however, no evidence has yet been obtained that the enzyme actually undergoes glycosylation. The protein has 18 cysteine residues, 6 of them clustered near the C-terminal portion of the protein. One of the cysteine residues is positioned only five residues away from the putative active site (see Discussion). This active site having the sequence His-Glu-X-X-His (residues 473–477) contains the two histidine residues, known to coordinate Zn^{2+} in several zinc metallopeptidases. The Glu residue in the active site is known to function in the bond-breaking process in several metallopeptidases. The -H-E-X-X-H- sequence seems to represent a unique motif common for zinc metallopeptidases (Jongeneel et al., 1989; Valee & Auld 1990) and is also present in thermolysin and related bacterial enzymes, as well as in a number of mammalian peptidases including endopeptidase 24.11 (enkephalinase), angiotensin converting enzyme, and several collagen-degrading enzymes (Devault et al., 1987; Soubrier et al., 1988; Goldberg et al., 1986; Collier et al., 1988).

Tissue Distribution of EP 24.15 mRNA. Northern blot hybridization analysis was performed with total RNAs from a variety of rat tissues in order to detect the distribution and size EP 24.15 mRNA. As shown in Figure 3, the analyses detected the presence of a mRNA with a length of approximately 2.6 kilobases only in testes. It is notable that other tissues had no detectable mRNA even after long exposures (data not shown), suggesting that there was at least a 20-fold lower concentration of EP 24.15 mRNA in nontestes tissues. However, significant enzymatic activities could be found in these tissues (Table I).

Discussion

Screening of a rat testes cDNA library led to the identification and isolation of a number of bacterial clones expressing EP 24.15. Some of the clones contained unusually high enzyme activities, making them an attractive source for isolation of large quantities of EP 24.15. In view of the great difficulties connected with the isolation of the enzyme from mammalian tissues, the bacterial clones might be useful in further studies on the relationship between the structure of the expressed fusion proteins and enzyme activity. The open reading frame of the cDNA encoding EP 24.15 is composed of 1935 nucleotides encoding a protein with 645 amino acid residues. A comparison of the amino acid sequence of the enzyme with that of thermolysin, a bacterial zinc metalloendopeptidase, and also with that of several mammalian zinc metallopeptidases including angiotensin converting enzyme (3.4.15.1) and enkephalinase (EC 3.4.24.11) failed to discern any homology beyond a short amino acid stretch at positions 470–479 that comprises the apparent catalytic site. Indeed, a search of the University of Geneva protein sequence data bank (SwissProt) and the Protein Identification Resource sequence data bank (PIR) revealed no significant homology with any known protein.

It is notable that while the Northern blot analysis detected abundant EP 24.15 mRNA in testes, no mRNA was detected in RNAs isolated from other tissues, including the brain. On the basis of the sensitivity of the Northern blot analytical procedure, the detection of 5% of the mRNA present in testes could have been expected. Since enzyme activities in brain and most other tissues were higher than 5% of those in testis, it is tempting to speculate that the half-life of the enzyme in most tissues is longer than in testes.

Like other members of the family of zinc-dependent metallopeptidases, with the exclusion of the carboxypeptidases, EP 24.15 contains the typical amino acid sequence at and around the active site that is represented by the motif X-X-X-H-E-X-X-H-X-X, in which the two histidine residues coordinate the Zn^{2+} in the active center and the glutamate is involved in the bond-breaking process. The three amino acid residues before the first histidine are generally represented by neutral amino acids, and the two residues after the second histidine are generally hydrophobic (Jongeneel et al., 1989). The sequence -T-Y-F-H-E-F-G-H-V-M- found in EP 24.15 conforms exactly with the above motif. The sequence -H-E-F-G-H- in the active center is the same as that found in human gelatinase; however, no homology was found with this enzyme beyond this five amino acid stretch.

Elucidation of the amino acid sequence of thermolysin (Titani et al., 1972) and X-ray diffraction analysis of its crystal structure (Matthews et al., 1972a,b) have shown that in addition to the Zn^{2+} coordinating 2 histidine residues (His-142 and His-146) in the active center, a third protein ligand is provided by Glu-166, separated from the second His residue in the active center by a 19 amino acid "spacer" (Vallee & Auld, 1990). Similar 19 amino acid spacers were also found in several related bacterial metalloendoproteinases (Valee & Auld, 1990). It is therefore of interest that a Glu residue separated from the His residues by a 19 amino acid spacer is also present in EP 24.15, although there is no evident homology between these enzymes, and also the sequence around this Glu residue is quite different from that in the bacterial proteases. It is therefore uncertain whether this Glu residue fulfills the same function in EP 24.15 as in the bacterial

enzymes, especially since no similarly spaced ligands are present in several mammalian zinc metalloendopeptidases such as endopeptidase 24.11 (Devault et al., 1987), collagen-degrading enzymes (Goldberg et al., 1986; Collier et al., 1988), and also in the angiotensin converting enzyme. This latter enzyme contains a glutamate residue (believed to be part of the active site) that is separated from the second histidine in the zinc motif by a spacer of 23 amino acid residues (Kumar et al., 1989). Further studies involving single amino acid mutations should elucidate the role of these residues in catalysis.

Endopeptidase 24.15 is inhibited by *p*-mercuribenzoate and by *N*-ethylmaleimide, two thiol blocking agents, although iodoacetamide, iodoacetic acid, and E-64, all cysteine proteinase inhibitors, have no effect on activity. EP 24.15 is apparently identical with Pz-peptidase (Barrett & Tisljar, 1989a) and probably also with oligoendopeptidase A (Camargo et al., 1973; Oliveira et al., 1976; Camargo et al., 1982). This last enzyme was claimed to be a cysteine oligoendopeptidase on the basis of inhibition by *p*-mercuribenzoate, partial inhibition by *N*-ethylmaleimide, and some activation by 2-mercaptoethanol, although dithiothreitol was shown to inhibit the enzyme (Orlowski et al., 1983). Because Pz-peptidase behaved in a similar manner, the thiol dependence of this enzyme was also stressed (Tisljar & Barrett, 1989b). The finding, however, that EP 24.15 is not inhibited by iodoacetamide, iodoacetic acid, and E-64, that it is inhibited by EDTA and *o*-phenanthroline, and that the activity of an enzyme inactivated by dialysis against EDTA can be fully restored by low concentrations of Zn^{2+} and also Co^{2+} clearly showed its metallopeptidase nature (Orlowski et al., 1983, 1989). This was also corroborated by the finding that substrate-related *H*-(carboxymethyl)peptides are potent specific inhibitors of EP 24.15 with K_i values in the nanomolar range (Chu & Orlowski, 1984; Orlowski et al., 1988). The presence in the cloned enzyme of an active site typical for a group of zinc metallopeptidases confirms the classification of the enzyme as a zinc metalloendopeptidase. The presence of a cysteine removed only five residues from the catalytic center can probably now account for some previously reported properties of the enzyme, such as activation by 2-mercaptoethanol and inhibition by *p*-mercuribenzoate and *N*-ethylmaleimide. It is possible that the lack of inhibition by iodoacetamide and iodoacetic acid could be explained on the basis that unlike with *p*-mercuribenzoate, the carboxymethylation of the cysteine residue by these small-size thiol blocking agents does not interfere with the access of substrates to the active site.

Specificity studies showed that EP 24.15 cleaves preferentially bonds on the carboxyl side of hydrophobic residues in the P_1 position whereas EP 24.11 cleaves bonds on the amino side of hydrophobic residues (position P_1'). Bacterial thermolysin contains a valine residue (valine-139) separated by a two amino acid spacer from the first histidine residue in the active center (Colman et al., 1972; Kester & Matthews, 1977). Similar valine residues are present in several related bacterial enzymes and also in endopeptidase 24.11 (valine-580), a mammalian enzyme with a thermolysin-like specificity. This valine residue is part of the hydrophobic binding pocket at the S_1' subsite of thermolysin that confers upon the enzyme its primary specificity directed toward bonds on the amino side of hydrophobic residues. In EP 24.15, the corresponding position is occupied by a Thr residue. Unlike EP 24.11, EP 24.15 does not prefer substrates with hydrophobic residues in the P_1' position (Orlowski et al., 1983, 1989). While the mode of substrate binding to the active site of EP 24.15 is still unknown, it is possible that the differences in the nature of the amino acids at these sites could contribute to the differences in substrate specificity between these enzymes.

Endopeptidase 24.15 is apparently involved in the metabolism of neuropeptides both in the central nervous system as well as in the periphery. The enzyme shows a rather high affinity toward such peptides as neurotensin, angiotensins I and II, dynorphin$^{1-8}$, α- and β-neoendorphin, and also several Met-enkephalin-containing peptides. Inhibitors of the enzyme have been shown to prolong the half-life in vivo of GnRH and thereby greatly increase the secretion of LH and FSH by the pituitary. Functional, structural and biochemical studies of the enzyme are accordingly of considerable interest for the physiology and pharmacology of the metabolism of bioactive peptides. Furthermore, specific inhibitors of the enzyme could constitute a new class of pharmacologically active agents that could modulate responses to these potent agents. The isolation of EP 24.15 from mammalian tissues presents considerable difficulties. The finding, therefore, that bacterial clones express large quantities of the active enzyme should be of importance for further studies on the function and biochemistry of the enzyme.

References

Acker, G. R., Molineaux, C., & Orlowski, M. (1987) *J. Neurochem. 48*, 284–292.

Barrett, A. J., & Tisljar, U. (1989a) *Biochem. J. 261*, 1047–1050.

Burnette, W. N. (1981) *Anal. Biochem. 112*, 195–203.

Camargo, A. C. M., Shapanka, R., & Greene, L. J. (1973) *Biochemistry 12*, 1838–1844.

Camargo, A. C. M., Fonseca, M. J. V., Caldo, H., & Carvalho, K. M. (1982) *J. Biol. Chem. 257*, 9265–9267.

Chu, T. G., & Orlowski, M. (1984) *Biochemistry 23*, 3598–3603.

Chu, T. G., & Orlowski, M. (1985) *Endocrinology 116*, 1418–1425.

Collier, I. E., Wilhelm, S. M., Eisen, A. Z., Marmer, B. L., Grant, G. A., Seltzer, J. L., Kronberger, A., He, C., Bauer, E. A., & Goldberg, G. I. (1988) *J. Biol. Chem. 263*, 6579–6587.

Colman, P. M., Jansonius, J. N., & Matthews, B. W. (1872) *J. Mol. Biol. 70*, 701–724.

Devault, A., Lazure, C., Nault, C., Le Moual, H., Seidah, N. G., Chretien, M., Kahn, P., Powell, J., Mallet, J., Beaumont, A., Roques, B. P., Crine, P., & Boileau, G. (1987) *EMBO J. 6*, 1317–1322.

Dixon, M. (1959) *Biochem. J. 55*, 170–171.

Goldberg, G. I., Wilhelm, S. M., Kronberger, A., Bauer, E. A., Grant, G. A., & Eisen, A. Z. (1986) *J. Biol. Chem. 261*, 6600–6605.

Jongeneel, C. V., Bouvier, J., & Bairoch, A. (1989) *FEBS Lett. 242*, 211–214.

Kerr, M. A., & Kenny, A. J. (1974) *Biochem. J. 137*, 477–488.

Kester, W. R., & Matthews, B. W. (1977) *Biochemistry 16*, 2506–2516.

Kumar, R. S., Kusari, J., Roy, S. N., Soffer, R. L., & Sen, G. C. (1989) *J. Biol. Chem. 264*, 16754–16758.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Lasdun, A., & Orlowski, M. (1990) *J. Pharmacol. Exp. Ther. 253*, 1265–1271.

Lasdun, A., Reznik, S., Molineaux, C. J., & Orlowski, M. (1989) *J. Pharmacol. Exp. Ther. 251*, 439–447.

Lowry, O. H., Rosebrough, A. L., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P., & Dupourque, D. (1972a) *Nature (London) 238*, 37–41.
Matthews, B. W., Colman, P. M., Jansonius, J. N., Titani, K., Walsh, K. A., & Neurath, H. (1972b) *Nature (London) 238*, 41–43.
Molineaux, C. J., Lasdun, A., Michaud, C., & Orlowski, M. (1988) *J. Neurochem. 51*, 624–633.
Morales, T. I., & Woessner, J. F., Jr. (1977) *J. Biol. Chem. 252*, 4855–4860.
Oliveira, E. B., Martins, A. R., & Camargo, A. C. M. (1976) *Biochemistry 15*, 1967–1974.
Orlowski, M., Michaud, C., & Chu, T. G. (1983) *Eur. J. Biochem. 135*, 81–88.
Orlowski, M., Michaud, C., & Molineaux, C. J. (1988) *Biochemistry 27*, 597–602.
Orlowski, M., Reznik, S., Ayala, J., & Pierotti, A. R. (1989) *Biochem. J. 261*, 951–958.
Pozsgay, M., Michaud, C., Liebman, M., & Orlowski, M. (1986) *Biochemistry 25*, 1292–1299.
Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G., & Corvol, P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9386–9390.
Stone, K. L., Lopresti, M. B., Crawford, J. M., DeAngelis, R., & Williams, K. R. (1989) *Protein sequence from microquantities of proteins and peptides* (Matsudaira, P., Ed.) Academic Press, New York.
Tisljar, U., & Barrett, A. J. (1989b) *Arch. Biochem. Biophys. 274*, 138–144.
Titani, K., Hermondson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972) *Nature (London) 238*, 35–37.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350–4354.
Valee, B. L., & Auld, D. S. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 220–224.

Polarity of Annealing and Structural Analysis of the RNase H Resistant α-5′-d[TACACA]·β-5′-r[AUGUGU] Hybrid Determined by High-Field ^{1}H, ^{13}C, and ^{31}P NMR Analysis†

William H. Gmeiner,‡ K. Ekambareswara Rao,‡ Bernard Rayner,§ Jean-Jacques Vasseur,§ François Morvan,§ Jean-Louis Imbach,§ and J. William Lown*,‡

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2, and Laboratoire de Chimie Bio-Organique, UA 488 CNRS, Université des Sciences et Techniques du Languedoc, Place E. Bataillon, 34060 Montpellier, Cedex 1, France

Received April 27, 1990; Revised Manuscript Received July 27, 1990

ABSTRACT: The novel hybrid duplex α-5′-d[TACACA]-3′·β-5′-r[AUGUGU]-3′ was analyzed extensively by 1D and 2D NMR methods. Two forms of the duplex exist in about an 80:20 ratio. Analysis of the exchangeable imino protons of the major component revealed that three AU and one AT base pair are present in addition to two GC base pairs, confirming that the duplex anneals in parallel orientation. The presence of the AT base pair, which can only be accounted for by a parallel duplex, was confirmed by a selective INEPT experiment, which correlated the thymidine imino proton to its C5 carbon. The lesser antiparallel form could be detected by exchangeable and nonexchangeable proton resonances in both strands. An exchange peak was observed in the NOESY spectrum for the thymidine methyl group resonance in both the predominant and lesser conformations, indicating the lifetime of the individual structures was on the millisecond time scale. The nonexchangeable protons of the predominant duplex were assigned by standard methods. The sugar pucker of the ribonucleosides was determined to be of the "S" type by a pseudorotation analysis according to Altona, with the *J*-couplings measured from the multiplet components of the phase-sensitive COSY experiment. The NOE pattern observed for the α-deoxynucleosides also suggested an S-type sugar pucker. The adoption of an S-type sugar pucker for both strands indicates that, in contrast to RNA·DNA duplexes formed exclusively from β-nucleotides, the α-DNA·β-RNA duplex may form a B-type helix. The ^{31}P resonances of the α and β strands have very different chemical shifts in the hybrid duplex and the difference persists above the helix melting temperature, indicating an intrinsic difference in ^{31}P chemical shift for nucleotides differing only in the configuration about the glycosidic bond.

The concept of using antisense oligonucleotides to block gene expression evolved from experiments carried out in the early 1980s. While it was recognized that prokaryotes employ an antisense mechanism to specifically control gene expression, the first demonstration that antisense oligonucleotides could block gene expression selectively in eukaryotic cells was given in studies by Izant and Weintraub (1984). The expression of a thymidine kinase gene injected into mouse fibroblast cells was significantly reduced if a plasmid that directed the production of antisense RNA for the thymidine kinase gene was coinjected.

† This investigation was supported by grants (to J.W.L.) from the Medical Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada and (to J.-L.I.) from CNRS and ARC (France).
* To whom correspondence should be addressed.
‡ University of Alberta.
§ Université des Sciences et Techniques du Languedoc.

0006-2960/90/0429-10329$02.50/0