Okada, Y., Nagase, H., & Harris, E. D. (1986) *J. Biol. Chem. 261*, 14245–14255.
Salo, T., Liotta, L. A., & Tryggvason, K. (1983) *J. Biol. Chem. 258*, 3058–3063.
Sanchez-Lopez, R., Nicholson, R., Gesnel, M. C., Matrisian, L. M., & Breathnach, R. (1988) *J. Biol. Chem. 263*, 11892–11899.
Sellers, A., Reynolds, J. J., & Meikle, M. C. (1978) *Biochem. J. 171*, 493–496.
Stricklin, G. P., Jeffrey, J. J., Roswit, W. T., & Eisen, A. Z. (1983) *Biochemistry 22*, 61–68.
Tryggvason, K., Höyhtyä, M., & Salo, T. (1987) *Biochim. Biophys. Acta 907*, 191–217.
Uhlén, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., & Lindberg, M. (1984) *J. Biol. Chem. 259*, 1695–1702.
Vater, C. A., Nagase, H., & Harris, E. D. (1983) *J. Biol. Chem. 258*, 9374–9382.
Vater, C. A., Nagase, H., & Harris, E. D. (1986) *Biochem. J. 237*, 853–858.
Welgus, H. G., & Stricklin, G. P. (1983) *J. Biol. Chem. 258*, 12259–12264.
Werb, Z., & Reynolds, J. J. (1975) *Biochem. J. 151*, 645–653.
Whitham, S. E., Murphy, G., Angel, P., Rahmsdorf, H. J., Smith, B. J., Lyons, A., Harris, T. J. R., Reynolds, J. J., Herrlich, P., & Docherty, A. J. P. (1986) *Biochem. J. 240*, 913–916.
Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., & Axel, R. (1979) *Cell 16*, 777–785.
Wilhelm, S. M., Collier, I. E., Kronberger, A., Eisen, A. Z., Marmer, B. L., Grant, G. A., Bauer, E. A., & Goldberg, G. I. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6725–6729.
Yamada, K. M. (1983) *Annu. Rev. Biochem. 52*, 761–799.

Expression of Two Kallikrein Gene Family Members in the Rat Prostate†

James M. Brady, Debora R. Wines, and Raymond J. MacDonald*
Department of Biochemistry, The University of Texas Southwestern Medical Center, Dallas, Texas 75235
Received November 21, 1988; Revised Manuscript Received February 22, 1989

ABSTRACT: We have characterized two kallikrein gene family members expressed in the prostate and submaxillary glands of rats. One mRNA (S3) is identical with the previously characterized submaxillary gland S3 mRNA that encodes an enzyme closely related to tonin. The second mRNA (P1) encodes a novel kallikrein-like enzyme that retains key amino acid residues responsible for the characteristic enzymatic cleavage specificity of kallikrein. Two P1-specific oligonucleotide probes derived from the P1 mRNA sequence were used to demonstrate the presence of P1 mRNA in the prostate and submaxillary glands and its absence in eight other rat tissues known to express one or more members of the kallikrein family. The P1-coding gene (rGK-8) was identified among genomic clones containing kallikrein family members by hybridization with a P1-specific oligonucleotide probe. The identification of the P1 gene was verified by nucleotide sequencing; the exon sequences of rGK-8 match the P1 mRNA sequence. The upstream region of rGK-8, where transcriptional regulatory elements likely reside, is very similar to that of other rat kallikrein family genes which are expressed in distinct tissue-specific patterns.

The kallikrein gene family encodes serine proteases that process biologically active peptides. This subfamily of serine proteases is characterized by limited substrate specificity and high sequence conservation when compared to other simple serine proteases. Tissue kallikrein is the best characterized enzyme of this family, which also includes tonin, the γ subunit of nerve growth factor (γ-NGF), epidermal growth factor binding proteins (EGF-BPs), prostate-specific antigen (PSA), and other less well-characterized proteases. Tissue kallikrein specifically cleaves the protein kininogen to release the potent vasodilatory peptide lysyl-bradykinin (Yamada & Erdos, 1982). Tonin cleaves angiotensinogen in vitro to release angiotensin II, a potent vasoconstrictor (Boucher et al., 1974). γ-NGF processes the precursor of nerve growth factor (Thomas et al., 1981), and the EGF-BPs process the precursor of epidermal growth factor (Lundgren et al., 1984; Drinkwater et al., 1987). Human PSA cleaves the high molecular weight seminal vesicle protein (Watt et al., 1986), causing liquification of the seminal fluid clot (Lilja et al., 1987). Thus, the biochemical role of the enzymes of the kallikrein family appears to be the selective cleavage of polypeptide precursors, principally prohormones [reviewed in Fuller and Funder (1986), Drinkwater et al. (1987), and MacDonald et al. (1988)].

In rodents, the kallikrein gene family comprises 12–24 members (Mason et al., 1983; Evans et al., 1987; Ashley & MacDonald, 1985a; Wines et al., 1989) disparately expressed in a variety of tissues. In the rat, kallikrein-like enzymes or their mRNAs have been detected in the submaxillary gland, sublingual and parotid glands, pancreas, prostate, kidney, spleen, pituitary gland, testis, brain, and liver (Tschesche et al., 1979; Swift et al., 1982; Ashley & MacDonald, 1985b; Pritchett & Roberts, 1987; Chao & Chao, 1987). For the mouse, from which all 24 kallikrein family genes have been characterized, only 1 gene encodes tissue (true) kallikrein (van Leeuwen et al., 1986); this is probably true for human, rat, and hamster as well. The other genes either encode proteins closely related to kallikrein or are pseudogenes (Evans et al.,

† This work was supported by Grant GM31689 from the National Institutes of Health. D.R.W. was supported by NIH Predoctoral Training Grant GM08203.
* Correspondence should be addressed to this author.

0006-2960/89/0428-5203$01.50/0

1987). The great similarity among family members makes it difficult to distinguish and identify individual genes, mRNAs, or enzymes. All genes and mRNAs of the family cross-hybridize to cDNA probes; the enzymes have similar catalytic activities and share common epitopes recognized by polyclonal antisera raised against individual family members. True kallikrein or its mRNA has been detected in the rat submaxillary gland, kidney, pancreas, parotid, and pituitary by using protein-specific monoclonal antibodies or gene-specific oligonucleotide probes (Chao & Chao, 1987; Ashley & MacDonald, 1985b; Pritchett & Roberts, 1987). Other members of the rat kallikrein gene family exhibit several unique patterns of tissue-specific expression.

Kallikrein-like enzymes in the prostate affect various aspects of reproductive physiology. Kallikrein or related family members increase testicular blood flow, accelerate spermatogenesis (Saito & Kumamoto, 1988), and process proteins important to male reproductive physiology (Watt et al., 1986). Prostate kallikrein-like enzymes may have an exocrine role as well; they are present in the prostatic fluid and seminal plasma (Fuller & Funder, 1986). They also may be involved in the maturation of nerve growth factor and epidermal growth factor, which are present at high concentrations in the prostate (Isackson et al., 1987).

The purpose of the present work is to identify the members of the rat kallikrein gene family that are expressed in the rat prostate. Characterization of the gene products may give insight into their physiological role. In addition, the analysis of the tissue-specific expression of individual members serves as a foundation for the study of regulatory variations within this family of duplicated genes that cause its members to be expressed at different levels in a variety of tissues. We have characterized two kallikrein-related mRNAs (designated P1 and S3) from rat prostate by cloning their cDNAs and determined their nucleotide sequences and the sequences of their encoded enzymes. The prostate S3 mRNA is identical with the rat submaxillary gland S3 mRNA described previously (Ashley & MacDonald, 1985a), which encodes an enzyme similar to tonin. The P1 mRNA is previously uncharacterized and encodes an enzyme similar to true kallikrein, but with unknown cleavage specificity and function. Hybridization with oligonucleotide probes specific for the P1 mRNA demonstrates that it is expressed selectively in the prostate and submaxillary glands. In addition, we have characterized the gene (rGK-8) encoding the P1 mRNA; the upstream region of this gene contains sequence elements common to other kallikrein-related genes.

Materials and Methods

Preparation and Screening of a Rat Prostate Double-Stranded Complementary DNA (ds-cDNA) Library. RNA was isolated from prostates of adult male Sprague-Dawley rats by the guanidine thiocyanate extraction procedure (Chirgwin et al., 1979; MacDonald et al., 1987). A rat prostate ds-cDNA library was constructed according to the method of Okayama and Berg (1982). The library was screened for kallikrein-like cDNA clones with a hybridization probe prepared from an M13mp8 subclone of the ds-cDNA insert of pcXP39 (Ashley & MacDonald, 1984) containing the 3′ 550 bp of rat pancreatic kallikrein mRNA (Swift et al., 1982), which encodes the carboxy-terminal 167 amino acids of true kallikrein plus the 3′-untranslated region.

The collection of kallikrein-related cDNA clones identified by hybridization with the cDNA was screened for individual family members by Southern blot hybridization with mRNA-specific oligonucleotide probes described under Results and in Table I. Oligonucleotides were synthesized by using phosphoramidite chemistry with an Applied Biosystems 380B automated DNA synthesizer. The oligonucleotide probes were end-labeled with [γ-^{32}P]ATP (NEN) and polynucleotide kinase (Maxam & Gilbert, 1980) and separated from γ-ATP by precipitation with cetylpyridinium bromide (Geck & Nasz, 1983). The selected cDNA clones were prehybridized overnight at 42 °C in oligonucleotide hybridization buffer [5 × SSC (1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 50 mM sodium phosphate, pH 7.0, 0.2% Ficoll 400, 0.2% bovine serum albumin, 0.2% polyvinylpyrrolidone, 250 μg/mL herring sperm DNA, and 1.0% SDS]. Hybridization was performed at 37 °C for 4–16 h under the conditions used for prehybridization with the addition of 1 × 10^6 cpm/mL labeled oligonucleotide. The filters were washed in three 100-mL changes, for 5 min each, of 6 × SSC and 0.05% sodium pyrophosphate, with 0.1% SDS added to the first wash, at a temperature 10 °C below the calculated dissociation temperature (T_d) of the oligonucleotide probe. T_d's were calculated from the empirical rules of Suggs et al. (1981): 4 °C for each G-C base pair and 2 °C for each A-T base pair.

Nucleotide Sequence Analysis. The nucleotide sequence of the 5′ end of the cloned P1 kallikrein-related cDNA was determined with the base-specific chemical cleavage procedures of Maxam and Gilbert (1980). Five sequencing reactions (G, dimethyl sulfate; G + A, formic acid; C + T, hydrazine; C, hydrazine + NaCl; A > C, NaOH) were employed to enhance sequence accuracy. The nucleotide sequence of the majority of the cloned P1 and all of the S3 kallikrein-related cDNA was determined by dideoxynucleotide sequencing (Sanger, 1977; Biggen et al., 1983). The cloned P1 and S3 cDNAs were digested with *Eco*RI restriction endonuclease and subcloned into M13mp19. Oligonucleotides complementary to known regions of the P1 and S3 cDNAs were used to select fragments for sequencing and as internal sequencing primers for dideoxynucleotide sequencing.

The sequencing schemes for the P1 cDNA and the P1 gene rGK-8 are summarized in Figure 1. The partial sequence of the P1 mRNA was obtained from subclones of recombinant plasmid pP1-12. Both strands were sequenced over 85% of the length of the ds-cDNA insert. The sequence of pP1-12 begins at codon −24 in the signal peptide region and extends through the poly(A) tail. Rescreening of the prostate cDNA library with an oligonucleotide specific for the P1 mRNA did not yield any recombinant plasmids with longer P1 cDNA inserts. The remainder of the P1 mRNA sequence was obtained from the λ genomic clone (see below) bearing the P1 gene rGK-8. More than 72% of the exon regions of the genomic clone was sequenced on both strands; more than 85% of the upstream sequence reported here was sequenced on both strands. The sequence of the prostate S3 kallikrein-related mRNA was obtained from recombinant plasmid pS3-4 identified in the prostate cDNA library. Both strands were sequenced over 80% of the length of the cDNA insert. The sequence begins with the 5′ noncoding region of the S3 mRNA and includes a poly(A) tail at the 3′ end.

Northern Blot Analysis of Glandular Kallikrein-Related mRNAs by Oligonucleotide Hybridization. Northern blots of polyadenylated RNA isolated from tissues of Sprague-Dawley rats of mixed sex were prepared as described previously (Ashley & MacDonald, 1985b). The Northern blots were hybridized with the P1-specific oligonucleotide probes $Ex2P1_{18}$ (dTTGCGGTTCATTAAAGAG) and $Ex3P1_{18}$ (dGTTCTTTATGATGTCCAG) derived from regions of the P1 mRNA with high sequence variation from other charac-

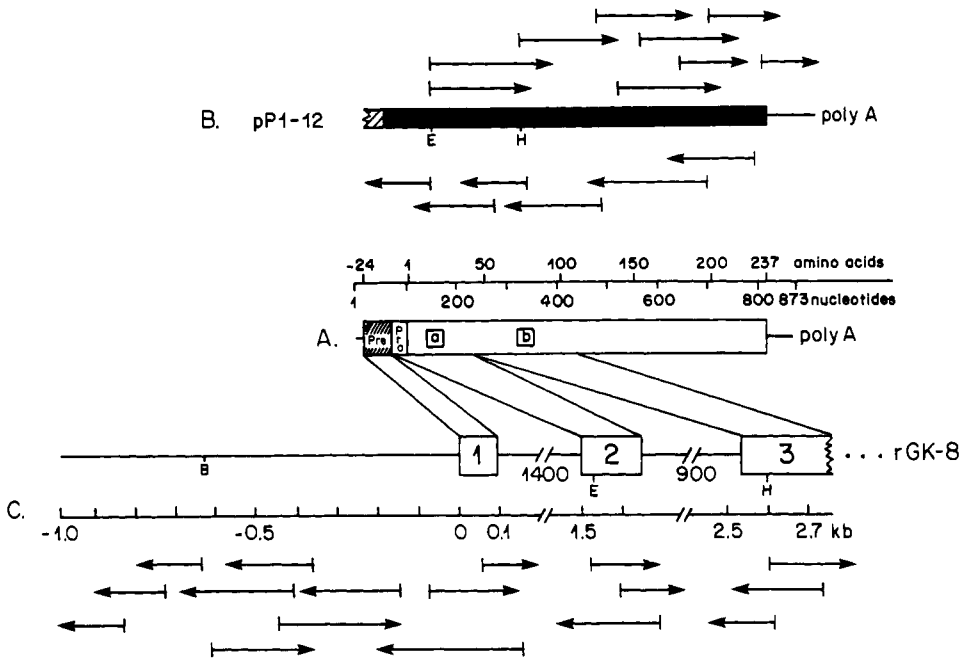

FIGURE 1: Organization and sequencing strategies for the rat P1 mRNA and rGK-8 gene. (A, center) The open horizontal rectangle delineates the amino acid coding region; the proposed signal peptide (Pre) and activation peptide (Pro) regions are indicated. The lines extending from the open rectangles represent the lengths of the noncoding regions of the mRNA; poly(A) is present at the 3′ end. The positions of the oligonucleotide probes used in the Northern hybridization studies are shown as boxes (a is $Ex2P1_{18}$; b is $Ex3P1_{18}$; also see Table I). (B) The extent of the cloned double-stranded cDNA pP1-12. The striped domain represents the region of the ds-cDNA cloning artifact. (C) The genomic clone rGK-8 is diagrammed; it contains the 5′ flanking sequence and exons 1, 2, and part of exon 3. Exons 4 and 5 plus the 3′ flanking genomic DNA remain uncloned. The λ vector DNA is indicated by three dots. The start, direction, and length of each sequencing run for the P1 cDNA and gene are indicated by the horizontal arrows. Restriction endonuclease sites are indicated by capital letters (E, *Eco*RI; H, *Hin*dIII; B, *Bgl*II).

terized kallikrein gene family members. The design of the gene-specific oligonucleotide probes is described under Results and in Table I. The hybridization and wash conditions were as described for the cDNA clone analysis. The stringency of the wash conditions was optimized by altering the temperature to eliminate cross-hybridization of the oligonucleotide probes to related sequences.

Identification and Analysis of the Gene Encoding the P1 mRNA. A genomic library of Sprague-Dawley rat liver DNA (Ashley & MacDonald, 1985b) in the λ vector L47.1 (Loenen & Brammer, 1980) was screened exhaustively with a long synthetic kallikrein probe to detect all related kallikrein family members (Wines et al., 1989). From this collection of cloned kallikrein family genes, a single genomic clone bearing the gene encoding P1 mRNA was identified initially by hybridization with the P1-specific oligonucleotide probe $Ex2P1_{18}$. The identity of the gene was verified by nucleotide sequencing of *Hin*dIII or *Eco*RI subclones in M13mp19 with the chain termination method (Sanger et al., 1977; Biggen et al., 1983) utilizing exon-specific oligonucleotides from regions conserved among characterized kallikrein family members and P1-specific oligonucleotides as primers (Yanisch-Perron et al., 1985).

Results

Isolation and Sequencing of Cloned ds-cDNA Sequences for Kallikrein-Related mRNAs of the Rat Prostate. Eleven cDNA clones containing kallikrein-related sequences were selected by screening a rat prostate cDNA library with a rat kallikrein cDNA probe (pcXP39; Swift et al., 1982) which cross-hybridizes with all known members of the rat kallikrein gene family. Southern analyses of the 11 prostate kallikrein cDNA plasmids with gene-specific oligonucleotide probes (Table I) identified cDNA clones representing characterized members of kallikrein gene family. The identification of individual family members is possible because of the presence of three short variant regions (nucleotide positions 168–189, 345–366, and 518–539 relative to the PS mRNA) in the previously characterized kallikrein genes and mRNAs (Ashley & MacDonald, 1985b). An oligonucleotide probe complementary to such a variant region of one mRNA hybridizes selectively to that mRNA (or cDNA clone). Kallikrein cDNA clones that do not hybridize to these oligonucleotide probes, therefore, are candidates for new members of the gene family.

The S3 kallikrein-related mRNA was previously detected in the rat prostate and submaxillary gland by Northern blot analysis (Ashley & MacDonald, 1985b). The presence of this S3 mRNA in the rat prostate was verified by the detection of two rat prostate cDNA clones that hybridize to an oligonucleotide probe, $Ex3S3_{21}$, that is specific for the S3 kallikrein mRNA. One nearly full-length S3 cDNA clone, pS3-4, from the rat prostate library was sequenced. The cDNA insert of pS3-4 is 859 nucleotides long, with a 5′-untranslated region of 36 nucleotides and a 3′-untranslated region of 46 nucleotides plus a poly(A) tail. The sequence of pS3-4 was identical with the S3 mRNA expressed in the rat submaxillary gland, demonstrating that the same structural gene is expressed in both tissues.

Another cDNA clone, pP1-12, hybridized to kallikrein cDNA but did not hybridize to oligonucleotides complementary to the other characterized kallikrein mRNAs. This novel kallikrein-related cDNA contained a nearly full-length mRNA sequence. The nucleotide sequence was obtained by sequencing the cDNA insert of recombinant plasmid pP1-12 (Figure 1); the partial P1 mRNA sequence contained in this cloned cDNA is 831 nucleotides long and extends from the second codon through the poly(A) tail.

The P1 kallikrein cDNA clone also contains an unusual cloning artifact that includes the deletion of some sequences and the duplication and inversion of other sequences at the 5′ end of the P1 mRNA (data not shown). Due to this cloning

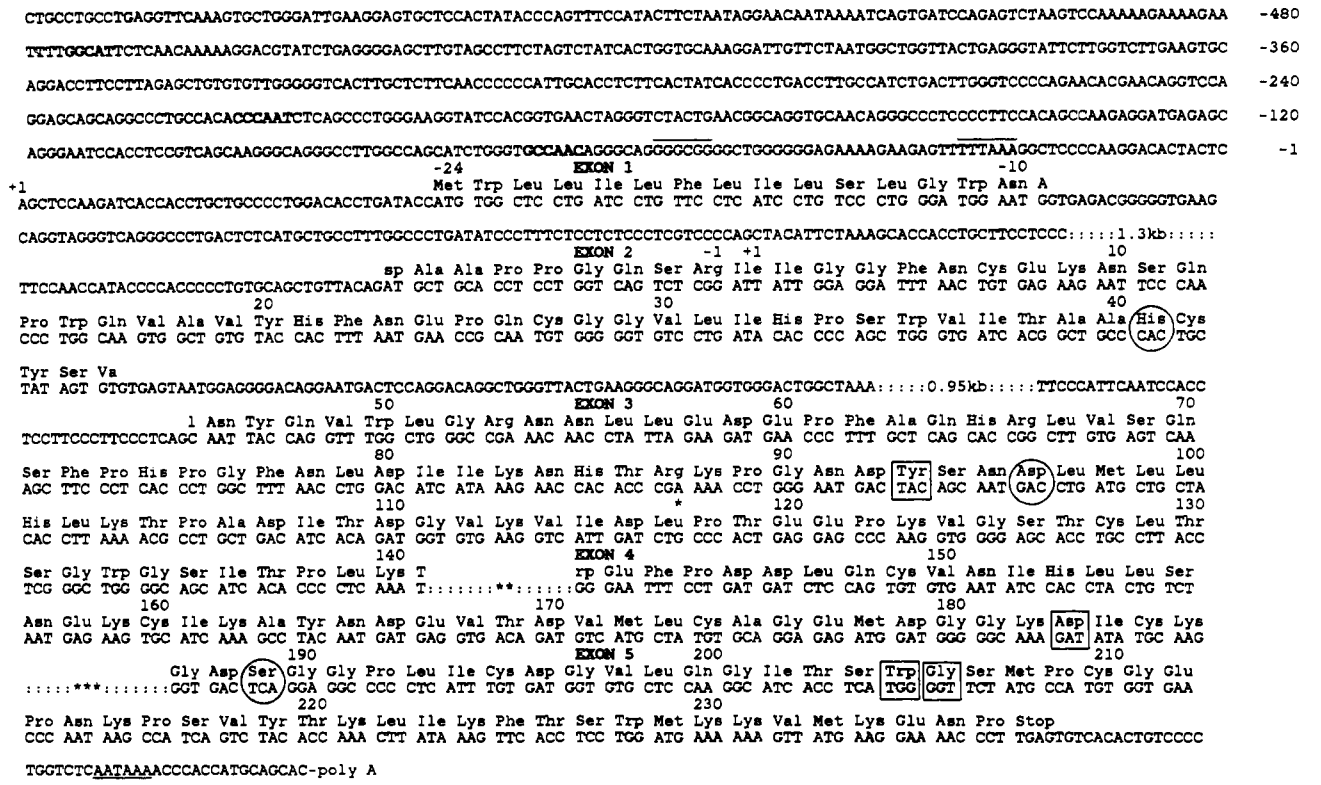

FIGURE 2: Nucleotide sequence and predicted protein product of the P1 mRNA and rGK-8 gene. The complete mRNA coding sequence was deduced from exons 1, 2, and 3 of the rGK-8 gene and the sequence of the pP1-12 cDNA spanning exonic regions 1 through 5 (see Figure 1). The nucleotide sequence of the cloned P1 kallikrein cDNA is identical with the sequence of rGK-8 from within the first codon (Met-24) through the end of rGK-8 at the codon for amino acid 117 (one asterisk). The lengths of intron 3 (two asterisks) and intron 4 (three asterisks) are not known; their presence and position are predicted from the conservation of exon/intron arrangements in this gene family for the rat, mouse, and human. The proposed TTTAAA box variant at −27/−22 and a potential Sp1 binding site at −57/−52 are *overlined*. The conserved AATAAA in the 3′-noncoding region of the mRNA is *underlined*. The estimated start site of transcription (+1) is based upon the 5′ end of other kallikrein family members. The amino acid residues of the serine protease catalytic triad are *circled*. Amino acid residues important in determining substrate specificity as discussed in the text are *boxed*.

Table I: Specific Oligonucleotide Probes Complementary to Distinctive Variant Regions of Kallikrein-Related mRNAs

oligomer designation[a]	no. of differences with respect to					
	PS	S1	S2	S3	K1[b]	P1
$Ex2P1_{18}$	9	*c*	9	12	11	–
$Ex3P1_{18}$	5	≥3[c]	8	6	7	–
$Ex3S2_{18}$	11	≥6[c]	–	5	7	8
$Ex3S3_{21}$	7	≥4[c]	5	–	4	10
$Ex3/4PS_{21}$	–	10	11	12	10	6
$Ex3/4S2_{21}$	11	3	–	3	10	11
$Ex3/4K1_{21}$	10	7	10	10	–	4

[a] The first part of the oligonucleotide designation gives the exon location of the variant region that the oligonucleotide is complementary to (e.g., Ex2 = exon 2). The second part indicates which kallikrein mRNA the oligonucleotide is specific for, and the subscript number indicates the length of the oligonucleotide. [b] The K1 mRNA sequence is a previously uncharacterized kallikrein-related mRNA expressed in the kidney (J. M. Brady and R. J. MacDonald, unpublished results). [c] The S1 mRNA sequence is not full-length, and its exon 2 sequence is not known. Only part of the exon 3 variable region is known for the S1 mRNA; the minimal number of differences with known mRNAs are shown.

artifact, the 5′ end of the P1 kallikrein mRNA is not correctly represented in pP1-12. Attempts to find additional cDNA clones that encode the 5′ end of this mRNA were unsuccessful. However, the correct sequence of the 5′ end of the P1 kallikrein-related mRNA was obtained from the sequence of the kallikrein family gene rGK-8, which encodes this mRNA (see below).

Isolation of Cloned Genomic Sequences for the P1 Gene. Recombinant λ clones bearing kallikrein family genes were identified by screening a rat genomic library with a long kallikrein probe that hybridizes to all gene family members (Wines et al., 1989). A P1-specific oligonucleotide, $Ex2P1_{18}$, (Table I) derived from a short subregion which is highly variable among the characterized kallikrein mRNAs was used to identify a single genomic clone bearing the P1-coding gene from among the collection of kallikrein family genomic clones. Sequence analysis confirmed that this gene, designated rGK-8, encodes the P1 mRNA. The genomic clone bearing rGK-8 contains exons 1 and 2 and the majority of exon 3; the 419 nucleotides of the mRNA sequence contained within the three exons is identical with the P1 cDNA, proving that rGK-8 encodes the P1 mRNA.

The correct sequence of the complete P1 mRNA and the derived amino acid sequence of the encoded enzyme are shown in Figure 2. The P1 kallikrein-related mRNA is estimated to be 873 nucleotides long with a 5′-untranslated region of 42 nucleotides, a 3′-untranslated region of 48 nucleotides, and a poly(A) tail. The P1 mRNA has 87%, 89%, 84%, and 83% nucleotide sequence identity with PS (true kallikrein), S1, S2 (tonin), and S3 kallikrein family mRNAs, respectively. The encoded protein of 265 amino acids contains a signal (pre) peptide sequence, an activation (pro) peptide sequence, and residues characteristic of kallikrein enzymes. Amino acid residues of the serine protease catalytic triad (histidine-41, aspartate-96, and serine-189) are present at the appropriate positions as are other amino acids characteristic of serine proteases. The P1 mRNA appears to code for a functional kallikrein-like enzyme with 80%, 83%, 75%, and 72% amino acid sequence identity with the enzymes encoded by PS, S1, S2, and S3 mRNAs, respectively. The high sequence identity and presence of key amino acid residues confirm that the P1

kallikrein-related mRNA is encoded by a member of the rat kallikrein gene family.

The 5′ flanking sequences of various kallikrein genes have been studied in detail because cis-acting transcriptional regulatory sequences are located in this region for several related serine protease genes (Boulet et al., 1986; Ornitz et al., 1985; Kruse et al., 1988). Comparison of the P1 gene (rGK-8) with the genes for true kallikrein (rGK-1) and tonin (rGK-2) (Wines et al., 1989) reveals common sequence elements that may be important for transcription promotion. The P1 gene contains a TATA box variant (TTTAAA) at −27/−22 (Figure 2), which is present in all of the characterized rat and mouse kallikrein family genes (Wines et al., 1989; Mason et al., 1983; Evans & Richards, 1985; van Leeuwen et al., 1986; Drinkwater et al., 1987). A potential Sp1 transcription factor binding site, GGGCGG, is present at −57/−52 in the P1 gene and is conserved in the rat true kallikrein and tonin genes. Sequence elements similar to the GCCAAT recognition motif of the CAT-box transcription factor (Santoro et al., 1988) are conserved in the true kallikrein, tonin, and P1 genes. These sites include GCCAAC/A at −70/−65, CCCAAT at −218/−213, and GCCAT at −475/−471 of the P1 gene. The three exons included in the genomic clone containing the P1 gene have the same intron boundaries as all other characterized rat (Wines et al., 1989; Chen et al., 1988), mouse (Mason et al., 1983; Evans et al., 1987), and human (Schedlich et al., 1987; Evans et al., 1988) kallikrein family genes.

Expression of P1 Kallikrein-Related mRNA in Other Tissues. Kallikrein-related mRNAs have previously been detected in the submaxillary gland, pancreas, prostate, spleen, kidney, testis, ovary, pituitary, liver, and parotid gland (Swift et al., 1982; Ashley & MacDonald, 1985b; Pritchett & Roberts, 1987; Fuller et al., 1985) by Northern blot analysis of polyadenylated RNA using a pancreatic kallikrein cDNA probe. Two different oligonucleotide probes ($Ex2P1_{18}$ and $Ex3P1_{18}$) were designed to determine which of these tissues contains P1 mRNA. The oligonucleotides are complementary to the P1 mRNA in 2 different variant regions within kallikrein mRNAs and contain 3–12 nucleotide differences (Table I) between P1 mRNA and the corresponding regions of the other 5 characterized rat kallikrein-related mRNAs. For those pairings with the minimum of three differences, the nucleotide mismatches are near the middle of the oligonucleotide. Thermal denaturation studies of such oligonucleotide/mRNA hybrids have shown that short oligonucleotides (18- to 21-mers) are specific for their cognate kallikrein-related mRNAs (J. M. Brady and R. J. MacDonald, unpublished results). The specificity of the two P1 oligonucleotides was shown by the lack of hybridization to the other cloned kallikrein cDNAs (Figure 3).

The P1-specific oligonucleotide probes $Ex2P1_{18}$ and $Ex3P1_{18}$ hybridize to a 0.9-kilobase (kb) RNA species in the submaxillary gland and prostate (Figure 3). In addition, at low stringency washes (6 × SSC, 42 °C), the probe $Ex3P1_{18}$ detected a 1.1-kb RNA at a low level in testis RNA (Figure 3B). With washes at higher stringency (6 × SSC, 55 °C), hybridization to the testis RNA disappears, whereas hybridization to the homologous submaxillary and prostate mRNAs does not decrease significantly (data not shown). The oligonucleotide probe $Ex2P1_{18}$, which hybridizes to submaxillary gland and prostate RNA, does not hybridize to testis RNA (Figure 3A). These results suggest the presence of an unidentified kallikrein-related mRNA in testis recognized by $Ex3P1_{18}$, but with mismatches that reduce the stability of the heterologous hybrid. The $Ex2P1_{18}$ oligonucleotide does not hybridize to testis mRNA; therefore, the unknown testis mRNA and the P1 mRNA contain significant nucleotide differences in the region this oligonucleotide covers. Similarly, a faint 2-kb pituitary RNA detected with $Ex3P1_{18}$ is lost at higher stringency washes and is not detected with the second ($Ex2P1_{18}$) oligonucleotide probe. These results demonstrate the utility of using two oligonucleotide probes complementary to two different regions for verifying sites of expression of members of highly related gene families.

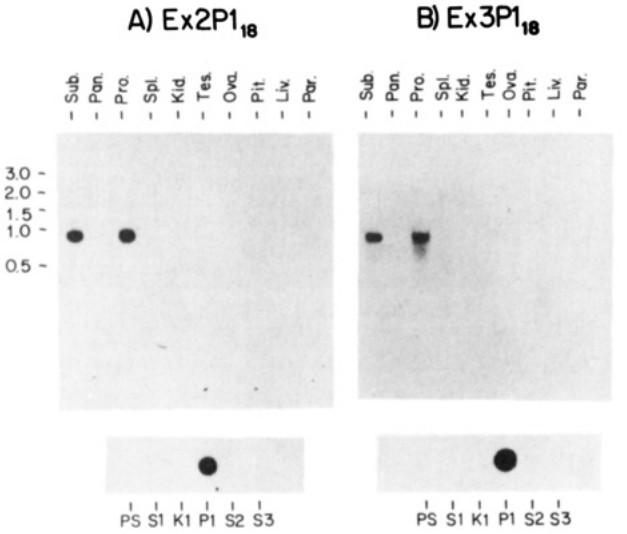

FIGURE 3: Northern blot analysis of the tissue distribution of P1 kallikrein-like mRNA. The lanes contain 1 μg of polyadenylated RNA (submaxillary gland) or 10 μg of polyadenylated RNA (pancreas, prostate, spleen, kidney, testis, ovary, pituitary, liver, and parotid). The size scale was derived from a 1-kb DNA ladder (BRL). The P1-specific oligonucleotide probes (A) $Ex2P1_{18}$ and (B) $Ex3P1_{18}$ are described in Table I. The lower portion of each panel shows the hybridization of control strips containing 25 ng of a cDNA plasmid for each of the characterized kallikrein mRNAs.

Discussion

Kallikrein-like enzymes in the prostate are believed to regulate blood flow (Saito & Kumamoto, 1988) and serve an important role in male reproductive physiology (Watt et al., 1986). We have cloned the mRNA sequences for two serine proteases that are at least partially responsible for the physiologically relevant kallikrein-like enzyme activities of the rat prostate gland. The encoded rat P1 enzyme retains key amino acids residues (Figure 4) at positions believed to be principal determinants of the correct cleavage of kininogen by true kallikrein to release vasoactive lysyl-bradykinin. The residues conserved between P1 and true kallikrein include aspartate-183 at the bottom of the binding pocket and glycine-206 at the mouth of the binding pocket (Bode et al., 1983). Aspartate-183 is the principal determinant of the characteristic kallikrein preference for cleavage after arginine residues in natural and synthetic substrates. The presence of Gly-206 is consistent with an open binding pocket to accommodate bulky amino acid side chains. Rat P1 and true kallikrein also contain tyrosine-93 and tryptophan-205, proposed to define the P_2[1] specificity of kallikrein. These residues form a hydrophobic sandwich (Bode et al., 1983) which traps hydrophobic amino acids in the P_2 substrate position and thereby determine sec-

[1] The nomenclature of Schechter and Berger (1967) is used to identify amino acid positions in the polypeptide substrate (P_n). Cleavage of the polypeptide substrate is at the carboxyl end of the P_1 residue; P_2 is the adjacent residue toward the amino terminus.

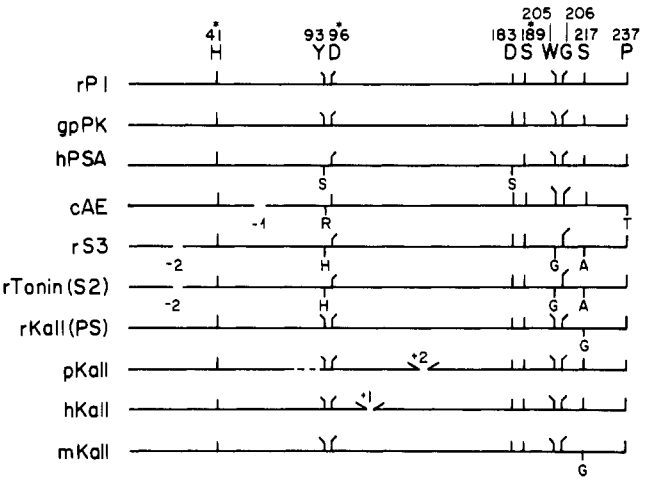

FIGURE 4: Comparison of key amino acid residues of the encoded P1 enzyme with true kallikrein and other kallikrein-like enzymes. The location and identity of key amino acid residues for serine protease catalytic activity (denoted by asterisks) and proposed for kallikrein-like preference for substrate cleavage (Chen & Bode, 1983) are indicated at the top. The presence of those residues in the other enzymes is indicated by lines extending above the horizontal lines. Key amino acid residues different from the P1-encoded enzyme are noted by letters below the horizontal lines. The rat kallikrein-like enzyme are denoted by the letter r; rP1 is the rat prostate enzyme encoded by rat prostate P1 kallikrein mRNA. The other enzymes compared are guinea pig prostate kallikrein-like enzyme (gpPK) (Dunbar & Bradshaw, 1987), human prostate-specific antigen (hPSA) (Watt et al., 1987; Lundwall & Lilja, 1987; Riegman et al., 1988), canine arginine esterase (cAE) (Chapdelaine et al., 1988), rat tonin-like enzyme (rS3), and rat tonin (rS2) (Ashley & MacDonald, 1985a). True kallikreins include those from rat (rKall) (Ashley & MacDonald, 1985a; Swift et al., 1982), porcine (pKall) (Bode et al., 1983), human (hKall) (Fukushima et al., 1985; Baker & Shine, 1985), and mouse (mKall) (mGK-6; van Leeuwen et al., 1986). The absence of one or two amino acid residues in cAE, rS3, and rS2 and the presence of one or two additional amino acid residues for hKall and pKall, respectively, are indicated. A polypeptide domain deleted by protease activity in pKall is shown by dashes.

ondary aspects of cleavage specificity. Other members of the family, such as rat tonin and canine arginine esterase, which do not retain these key amino acid residues (see Figure 4) also do not retain the cleavage specificity for kininogen characteristic of kallikrein. There is a difference between P1 and true kallikrein at the critical amino acid residue 217, which by analogy to porcine kallikrein (Bode et al., 1983) resides at the mouth of the binding pocket. The P1 enzyme has a serine at residue 217, while rat true kallikrein has a glycine at this position. However, this residue is variable among true kallikreins from various species; porcine and human true kallikreins contain Ser-217 while rat and mouse true kallikreins contain Gly-217, yet all maintain the same substrate specificity. Either glycine or serine at this position would be expected to maintain an open binding pocket to accommodate bulky amino acid side chains. The overall sequence conservation and the presence of similar key amino acid residues between true kallikrein and P1 suggest that the P1 enzyme may perform a biochemical function similar to that of true kallikrein.

Guinea pig prostate kallikrein (gpPK), the major arginine esteropeptidase in guinea pig prostate (Dunbar & Bradshaw, 1987), retains key amino acid residues identical with the rat P1 enzyme and the true kallikreins from porcine and human (see Figure 4). These include Asp-183, Gly-206, and Ser-217, the principal determinants of an open binding pocket and cleavage after basic amino acids. Guinea pig PK also retains Tyr-93 and Trp-205, which contribute to the kallikrein preference for hydrophobic amino acids in the P_2 substrate position. However, the substrate cleavage specificity and expression pattern of gpPK have not been reported, and the biochemical and physiological roles of gpPK remain unknown. Although NGF and EGF are present at high concentrations in the prostate (Harper et al., 1982), gpPK does not form a stable complex with β-NGF or EGF and does not appear to be involved in processing NGF or EGF (Isackson et al., 1987). As proposed for P1 in the rat, gpPK may function as an alternative true kallikrein in the guinea pig. The presence of an alternative kallikrein may be a common mechanism in rodents for the expression of a kinin-producing activity in the prostate gland.

Another biological role of kallikrein-like enzymes in the prostate is exemplified by a human prostatic secretory enzyme, prostate-specific antigen (PSA) (Watt et al., 1986). Human PSA cleaves high molecular weight seminal vesicle protein to a number of low molecular weight proteins responsible for the liquification of the seminal fluid clot (Lilja et al., 1987). It is unlikely that the P1 enzyme is the rat counterpart of hPSA, however, because differences in critical amino acid residues (Ser for Tyr-93; Ser for Asp-183) suggest very different peptide cleavage specificities. Similarly, differences in key amino acid residues between the P1 enzyme and canine arginine esterase (cAE) (Chapdelaine et al., 1988), which is one of the major proteins in dog prostate (Lazure et al., 1984) and seminal plasma (Isaacs, 1984) and lacks kininogenase activity (Frenette et al., 1985), suggest that the P1 and cAE enzymes perform different functions.

The second kallikrein-related mRNA (S3) characterized in the rat prostate encodes a protein that is very closely related to rat tonin (Ashley & MacDonald, 1985b). The S3 enzyme retains the amino acid residues aspartate-183, alanine-217, histidine-93, and glycine-205 that determine much of the size, shape, and ionic character of the substrate binding pocket of tonin (Fujinaga & James, 1987). In comparison to true kallikrein, these few changes at critical residues plus altered conformations of polypeptide loops near the substrate binding site (Fujinaga & James, 1987) may account for tonin's altered substrate cleavage properties and its specific cleavage of angiotensinogen. Because of the close similarities, the S3 enzyme may carry out an enzymatic function in the prostate similar to that of tonin, but the actual biological substrate of either enzyme remains to be determined.

The tissue-specific expression pattern of the rat P1 kallikrein-related mRNA has been determined by using two oligonucleotide probes designed to detect P1 mRNA specifically. Of the 10 rat tissues known to express 1 or more kallikrein family members, P1 mRNA was detected solely in the prostate and submaxillary glands. The specificity of the two P1-specific oligonucleotide probes was verified by their hybridization to P1 cDNA, but not to cDNAs of the other five characterized kallikrein family mRNAs. Although this experiment confirms the specificity of the P1 oligonucleotides relative to the known kallikrein family members, cross-hybridization with the mRNAs for other uncharacterized family members cannot be evaluated experimentally at this time. However, the routine use of two oligonucleotide probes derived from separate mRNA variant regions helps ensure specificity. Moreover, the restricted pattern of expression of P1 mRNA in only the prostate and submaxillary glands indicates that the P1-specific probes do not detect any of the kallikrein family mRNAs expressed in the other tissues examined.

The S3 kallikrein-related mRNA is expressed in the rat prostate at a high level and in the submaxillary gland at a moderate level (Wines et al., 1989). This expression pattern agrees with previous studies using a second oligonucleotide

specific for S3 mRNA (Ashley & MacDonald, 1985b; Clements et al., 1988). We had reported previously (Ashley & MacDonald, 1985b) that tonin (S2) mRNA was present in the prostate as well, based on hybridization with a single oligonucleotide probe. The oligonucleotide probe used in that study, however, cross-hybridizes with S3 mRNA. The use of a second oligonucleotide probe for S2 mRNA complementary to a more divergent mRNA region confirmed the presence of tonin mRNA in the submaxillary gland, and its absence in the prostate gland (Wines et al., 1989). These results illustrate the advantage of using two independent mRNA-specific oligonucleotides. Therefore, P1 and S3 are the only characterized kallikrein-related mRNAs expressed in the rat prostate.

The characterization of kallikrein-related mRNAs has facilitated the design of oligonucleotide probes that identify active kallikrein genes and the tissue-specific expression of their mRNAs. This approach has been used to identify a genomic DNA clone containing the kallikrein family gene rGK-8, which encodes the P1 mRNA expressed in the submaxillary gland and prostate. Further study of rGK-8 and other active kallikrein genes should reveal the general nature of tissue-specific regulatory elements as well as variations within a multigene family that direct the selective expression of kallikrein-related genes to different tissues.

Acknowledgments

We thank Dr. Patricia Ashley for construction of the prostate cDNA library, Lily Yamasaki for prescreening cDNA clones, Helen Aronovich for synthesis of the oligonucleotides utilized in this study, and Dr. Galvin Swift for critical reading of the manuscript.

Registry No. DNA (rat prostate gland kallikrein P1 mRNA complementary), 120665-53-4; prekallikreinogen (rat clone pP1-12 reduced), 120665-55-6; kallikrein (rat clone pP1-12 reduced), 120665-54-5; kallikrein, 9001-01-8; kallikrein S3, 9055-02-1; kallikreinogen (rat clone pP1-12 reduced), 120665-56-7; RNA (rat clone pP1-12 kallikrein-specifying messenger), 120665-57-8; prekallikreinogen, 84628-83-1.

References

Ashley, P. L., & MacDonald, R. J. (1984) *Anal. Biochem. 140*, 95–103.

Ashley, P. L., & MacDonald, R. J. (1985a) *Biochemistry 24*, 4512–4520.

Ashley, P. L., & MacDonald, R. J. (1985b) *Biochemistry 24*, 4520–4527.

Baker, A. R., & Shine, J. (1985) *DNA 4*, 445–450.

Biggen, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237–282.

Boucher, R., Asselin, J., & Genest, J. (1974) *Circ. Res. 34 (Suppl. I)*, 203–209.

Boulet, A. M., Erwin, C. R., & Rutter, W. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3599–3603.

Chao, J. (1983) *J. Biol. Chem. 258*, 4434–4439.

Chao, J., & Chao, L. (1987) *Biochim. Biophys. Acta 910*, 233–239.

Chapdelaine, P., Ho-Kim, M. Y., Tremblay, R. R., & Dube, J. Y. (1988) *FEBS Lett. 232*, 187–192.

Chen, Y. P., Chao, J., & Chao, L. (1988) *Biochemistry 27*, 7189–7196.

Chen, Z., & Bode, W. (1983) *J. Mol. Biol. 164*, 283–311.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Clements, J. A., Matheson, B. A., Wines, D. R., Brady, J. M., MacDonald, R. J., & Funder, J. W. (1988) *J. Biol. Chem. 263*, 16132–16137.

Drinkwater, C. C., Evans, B. A., & Richards, R. I. (1987) *Biochemistry 26*, 6750–6756.

Dunbar, J. C., & Bradshaw, R. A. (1987) *Biochemistry 26*, 3471–3478.

Evans, B. A., & Richards, R. I. (1985) *EMBO 4*, 133–138.

Evans, B. A., Drinkwater, C. C., & Richards, R. I. (1987) *J. Biol. Chem. 262*, 8027–8034.

Evans, B. A., Yun, Z. X., Close, J. A., Tregear, G. W., Kitamura, N., Nakanishi, S., Callen, D. F., Baker, E., Hyland, V. J., Sutherland, G. R., & Richards, R. I. (1988) *Biochemistry 27*, 3124–3129.

Frenette, G., Dube, J. Y., & Tremblay, R. R. (1985) *Biochim. Biophys. Acta 838*, 270–276.

Fujinaga, M., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 373–396.

Fukushima, D., Kitamura, N., & Nakanishi, S. (1985) *Biochemistry 24*, 8037–8043.

Fuller, P. J., & Funder, J. W. (1986) *Kidney Int. 29*, 953–964.

Fuller, P. J., Clements, J. A., Whitfield, P. L., & Funder, J. W. (1985) *Mol. Cell. Endocrinol. 39*, 99–105.

Geck, P., & Nasz, I. (1983) *Anal. Biochem. 135*, 264–268.

Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.

Harper, G. P., Glanville, R. W., & Thoenen, H. (1982) *J. Biol. Chem. 257*, 8541–8547.

Isaacs, W. B., & Coffey, D. S. (1984) *J. Biol. Chem. 259*, 11520–11526.

Isackson, P. J., Dunbar, J. C., & Bradshaw, R. A. (1987) *J. Cell. Biochem. 33*, 65–75.

Kruse, F., Komro, C. T., Michnoff, C. H., & MacDonald, R. J. (1988) *Mol. Cell. Biol. 8*, 893–902.

Lazure, C., Leduc, R., Seidah, N. G., Chretien, M., Dube, J. Y., Chapdelaine, P., Frenette, G., Paquin, R., & Tremblay, R. R. (1984) *FEBS Lett. 175*, 1–6.

Lilja, H., Oldbring, J., Rannevik, G., & Laurell, C. B. (1987) *J. Clin. Invest. 80*, 281–285.

Loenen, W. A. M., & Brammer, W. J. (1980) *Gene 10*, 249–259.

Lundgren, S., Ronne, H., Rask, L., & Peterson, P. A. (1984) *J. Biol. Chem. 259*, 7780–7784.

Lundwall, A., & Lilja, H. (1987) *FEBS Lett. 214*, 317–322.

MacDonald, R. J., Swift, G. H., Przybyla, A. E., & Chirgwin, J. M. (1987) *Methods Enzymol. 152*, 219–227.

MacDonald, R. J., Margolius, H. S., & Erdos, E. B. (1988) *Biochem. J. 253*, 313–321.

Mason, A. J., Evans, B. A., Cox, D. R., Shine, J., & Richards, R. I. (1983) *Nature (London) 303*, 300–307.

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

Okayama, H., & Berg, P. (1982) *Mol. Cell. Biol. 2*, 161–170.

Ornitz, D. M., Palmiter, R. D., Hammer, R. E., Brinster, R. L., Swift, G. H., & MacDonald, R. J. (1985) *Nature (London) 313*, 600–602.

Pritchett, D. B., & Roberts, J. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5529–5549.

Riegman, P. H. J., Klaassen, P., van der Korput, J. A. G. M., Romijn, J. C., & Trapman, J. (1988) *Biochem. Biophys. Res. Commun. 155*, 181–188.

Saito, S., & Kumamoto, Y. (1988) *Arch. Androl. 20*, 51–65.

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

Santoro, C., Mermod, N., Andrew, P. C., & Tijan, R. (1988) *Nature (London) 334*, 218–223.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157-162.
Schedlich, L. J., Bennetts, B. H., & Morris, B. J. (1987) *DNA 6*, 429-437.
Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K., & Wallace, B. (1981) in *Developmental Biology Using Purified Genes* (Brown, D. D., Ed.) pp 683-693, Academic Press, New York.
Swift, G. H., Dagorn, J.-C., Ashley, P. L., Cummings, S. W., & MacDonald, R. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7263-7267.
Thomas, K. A., Baglan, N. C., & Bradshaw, R. A. (1981) *J. Biol. Chem. 256*, 9156-9166.
Tschesche, H., Mair, G., Godec, G., Fiedler, F., Ehret, W., Hirschauer, C., Lemon, M., & Fritz, H. (1979) *Adv. Exp. Med. Biol. 120A*, 245-260.
van Leeuwen, B. H., Evans, B. A., Tregar, G. W., & Richards, R. I. (1986) *J. Biol. Chem. 261*, 5529-5535.
Watt, K. W. K., Lee, P.-J., M'Timkulu, T., Chan, W., & Loor, R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3166-3170.
Wines, D. R., Brady, J. M., & MacDonald, R. J. (1989) *J. Biol. Chem.* (in press).
Yamada, K., & Erdos, E. G. (1982) *Kidney Int. 22*, 331-337.
Yanisch-Perron, C., Viera, J., & Messing, J. (1985) *Gene 33*, 103-119.

Terminator-Distal Sequences Determine the in Vitro Efficiency of the Early Terminators of Bacteriophages T3 and T7†

Alice Telesnitsky and Michael J. Chamberlin*
Department of Biochemistry, University of California, Berkeley, California 94720
Received November 29, 1988; Revised Manuscript Received February 27, 1989

ABSTRACT: Bacteriophages T3 and T7 contain homologous terminators for *Escherichia coli* RNA polymerase that restrict early phage transcription to the leftmost 20% of the linear phage genomes. These two terminators serve equally well as ρ-independent terminators in vivo, but their in vitro efficiencies and sensitivity to salt and nucleotide concentrations differ dramatically. Sequence analysis shows that the T7 and T3 terminators differ at only two sites in the region normally accepted as defining terminator function. In order to determine which structural features of these two terminators are responsible for their functional differences, a series of hybrid terminators were constructed in which structural features of the two terminators were systematically interchanged. Transcription of hybrid terminator templates revealed that sequences downstream of the termination release sites are responsible for the differences in efficiency of in vitro termination. These sequences also determine the sensitivity of these terminators to elevated salt concentrations and to alterations of substrate concentrations. Alteration of the sequences in the region between three and seven nucleotides downstream of the final T7Te release site is sufficient to reduce termination efficiency to that of T3Te, and point mutations in this region yield terminators with intermediate efficiency. Hence, the determinants of ρ-independent terminator efficiency in vitro must include elements of the transcription complex other than the structure of the 3′ end of the transcript. The termination differences between T7Te, T3Te, and their hybrid derivatives are overcome in vivo; all of these sites become very efficient. This finding further supports the hypothesis that protein factors or other cellular features enhance the efficiency and specificity of ρ-independent terminators in vivo.

Transcriptional terminators for bacterial RNA polymerases have generally been divided into two classes—ρ-dependent and ρ-independent—on the basis of their response to the termination factor, ρ. ρ-independent termination sites are defined as those that can block transcription in vitro in the absence of factors other than RNA polymerase. These terminators often consist of a GC-rich region of dyad symmetry followed by a T-rich region in the coding strand which immediately precedes the transcriptional stop site (Adhya & Gottesman, 1978; Brendel & Trifanov, 1984, 1986). Numerous studies involving base analogues (Farnham & Platt, 1980, 1981), heteroduplexes (Ryan & Chamberlin, 1983), and mutant terminators (Yanofsky, 1981) support the model that sequences 40-50 base pairs upstream of the 3′ ends of transcripts are essential elements of ρ-independent terminators. A widely accepted model for ρ-independent termination (Adhya & Gottesman, 1978; Yanofsky, 1981) postulates that termination occurs after the GC-rich region forms a stable base-paired stem-loop structure in the RNA which causes the polymerase to pause. It is supposed that the elongating polymerase and transcript are then released due to the instability of the rU·dA base pairs which often comprise the remainder of the transcript/template hybrid (Adhya & Gottesman, 1978; Martin & Tinoco, 1980). By this model, there are two critical sequences that define the terminator: the dyad symmetry which encodes the RNA stem-loop, and the 3′ tail sequence between the base of the stem and the 3′ end of the RNA that determines the termination release site.

The early terminators of the bacteriophages T7 and T3 have very similar sequences (Briat et al., 1987; Figure 1). Both can be classified as ρ-independent terminators since they have typical ρ-independent terminator stem-loop and tail structures,

† This research was supported by a research grant (GM 12010) from the National Institute of General Medical Sciences.
* To whom correspondence should be addressed.

0006-2960/89/0428-5210$01.50/0