Johnson, P., & Smillie, L. B. (1975) *Biochem. Biophys. Res. Commun. 64*, 1316-1322.
Kress, M., Huxley, H. E., Faruqi, A. R., & Hendrix, J. (1986) *J. Mol. Biol. 188*, 325-342.
Lamkin, M., Tao, T., & Lehrer, S. S. (1983) *Biochemistry 22*, 3053-3058.
Leavis, P. C., & Gergely, J. (1984) *CRC Crit. Rev. Biochem. 16*, 235-305.
Lehrer, S. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3377-3381.
Lehrer, S. S. (1978) *J. Mol. Biol. 118*, 209-226.
Lehrer, S. S., & Kerwar, G. (1972) *Biochemistry 11*, 1211-1217.
Lehrer, S. S., & Morris, E. P. (1982) *J. Biol. Chem. 257*, 8073-8080.
Macgregor, R. B., & Weber, G. (1986) *Nature (London) 319*, 70-73.
Mak, A., Lewis, W., & Smillie, L. (1979) *FEBS Lett. 105*, 232-234.
Marriott, G. J. (1987) Ph.D. Thesis, University of Illinois.
McLachlan, A. D., & Stewart, M. (1978) *J. Mol. Biol. 103*, 271-298.
Morris, E. P., & Lehrer, S. S. (1984) *Biochemistry 23*, 2214-2220.
Nagashima, H., & Asakura, S. (1982) *J. Mol. Biol. 155*, 409-428.
O'Brien, E. J., Gillis, J. M., & Couch, J. (1975) *J. Mol. Biol. 99*, 461-475.
Okamoto, Y., & Sekine, T. (1985) *J. Biochem. (Tokyo) 98*, 1143-1145.
Phillips, G., Fillers, J. P., & Cohen, C. (1986) *J. Mol. Biol. 192*, 111-131.
Potekhin, S. A., & Privalov, P. L. (1982) *J. Mol. Biol. 159*, 519-535.
Potter, J. D., & Gergely, J. (1974) *Biochemistry 13*, 2697-2703.
Prendergast, F. G., Meyer, M., Carlson, G. L., Iida, S., & Potter, J. D. (1983) *J. Biol. Chem. 258*, 7541-7554.
Ruiz-Opazo, N., & Nadal-Ginard, B. (1987) *J. Biol. Chem. 262*, 4755-4765.
Smillie, L. B. (1979) *Trends Biochem. Sci. (Pers. Ed.) 4*, 151-155.
Stewart, M. (1975) *FEBS Lett. 53*, 5-7.
Tao, T., & Lamkin, M. (1984) *FEBS Lett. 168*, 169.
Tao, T., Lamkin, M., & Lehrer, S. S. (1983) *Biochemistry 22*, 3059-3066.
Tao, T., Gong, B.-J., & Leavis, P. C. (1987) *Biophys. J. 51*, 27a.
Weber, G., & Farris, F. J. (1979) *Biochemistry 18*, 3075-3078.
Weeds, A. G., & Pope, B. (1977) *J. Mol. Biol. 111*, 129-157.
Williams, D. L., & Swenson, C. A. (1981) *Biochemistry 20*, 3856-3864.
Williams, D. L., & Greene, L. E. (1983) *Biochemistry 22*, 2270-2274.
Woods, E. F. (1976) *Aust. J. Biol. Sci. 29*, 405-418.

Multiple Species of Myeloperoxidase Messenger RNAs Produced by Alternative Splicing and Differential Polyadenylation[†,‡]

Kazuya Hashinaka, Chika Nishio, Sook-Jin Hur, Fumio Sakiyama, Susumu Tsunasawa, and Michiyuki Yamada*
Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565, Japan
Received January 20, 1988; Revised Manuscript Received March 24, 1988

ABSTRACT: Three clones of full-length cDNA encoding human myeloperoxidase were isolated from a human leukemia HL-60 cell cDNA library in λgt10 and characterized. Analysis of the nucleotide sequence of one of the cDNA clones, λMP-H17, indicated that the cDNA contained 3207 bp with an open reading frame of 2238 bp, a 5′ noncoding region of 159 bp, a 3′ noncoding region of 800 bp, and a poly(A) tail of 10 bp. cDNA of the two other clones, λMP-H7 and λMP-H14, each contained insertions with shorter sequences of 96 and 82 bp, respectively, on the open reading frame of λMP-H17 cDNA. A myeloperoxidase genomic clone was isolated, and the structure of its 5′ region was determined and compared with the structures of these cDNAs. The comparison revealed that the three cDNAs were derived from myeloperoxidase mRNAs produced by alternative splicing from a transcript of the single gene. Nucleotide sequence analysis of the 3′ region of the cDNAs of several clones indicated that the mRNAs were polyadenylated at five different sites. Amino acid sequence determination of the amino-terminal and carboxy-terminal portions of the myeloperoxidase light and heavy chains revealed that, during processing of a precursor polypeptide into the mature protein, the amino-terminal polypeptide, the small peptide between the light and heavy chains, and the carboxy-terminal amino acid were excised.

Myeloperoxidase is a major hemoprotein present in azurophilic granules of polymorphonuclear leukocytes and is associated with a chloride ion mediated bacteriocidal action of these cells and with modulation of the metabolites generated in response to inflammation (Klebanoff, 1975; Henderson & Klebanoff, 1983). Multiple forms of myeloperoxidase were isolated from human myeloid leukemia HL-60 cells (Yamada et al., 1981) and human leukocytes (Pember et al., 1983; Morita et al., 1986; Miyasaki et al., 1986). The enzymes consisted of two molecules of a 59 000-Da[1] chain and two

[†] This work was partly supported by Grants-in-Aid for Cancer Research from the Ministry of Education, Science and Culture of Japan.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02831.
* To whom correspondence should be addressed.

0006-2960/88/0427-5906$01.50/0

molecules of a 15 000-Da chain (Andrews & Krinsky, 1981; Yamada et al., 1981). Slight differences in the amino acid compositions of these enzymes were reported (Pember et al., 1983; Morita et al., 1986). However, details of the heterogeneity of the proteins are not yet known.

Recently, Morishita et al. (1987a) and Johnson et al. (1987) isolated a full-length cDNA encoding human myeloperoxidase and determined its nucleotide sequence. These two cDNAs had the same nucleotide sequence, except that only one of them had a 3′ long untranslated region. Therefore, the amino acid sequence deduced from the two cDNAs was the same and included the light and heavy chains of myeloperoxidase. We isolated a cDNA encoding the carboxyl portion of myeloperoxidase (Yamada et al., 1987). The nucleotide sequence of the cDNA was included in that of one of the full-length cDNAs. In vitro translation of the myeloperoxidase mRNA selected by cDNA hybridization also indicated that myeloperoxidase is synthesized as a single-chain precursor (Yamada et al., 1987).

As an initial step in elucidating the heterogeneity of myeloperoxidase, we cloned and characterized various forms of full-length cDNAs encoding the protein. This work indicated that heterogeneous mRNAs were formed by alternative splicing and different polyadenylations. The amino-terminal and carboxy-terminal portions of the light chain and the heavy chain of myeloperoxidase were also determined. The results are discussed in relation to the synthesis of myeloperoxidase.

Experimental Procedures

Chemicals. All radiochemicals and a sequencing kit were purchased from Amersham; AMV reverse transcriptase was from Life Science; human placenta RNase inhibitor and RNase H were from Wako Pure Chemicals; DNA polymerase I, Klenow enzyme, and T4 DNA polymerase were from Toyobo; *Escherichia coli* DNA ligase and *Eco*RI methylase were from New England Biologicals; *Eco*RI-digested λgt10 DNA and Packergene were from Promega Biotec; a synthetic oligodeoxynucleotide was made by Takara Shuzo Co. Ltd.; carboxypeptidase Y was from Carbiotech; carboxypeptidase A (type I DFP) and carboxypeptidase B (DFP) were from Sigma Chemical Co.

Construction and Screening of a cDNA Library. Poly(A+) RNA from HL-60 cells was isolated by a modification of the guanidine thiocyanate method (Kasugai & Yamada, 1986) and then was enriched in myeloperoxidase mRNA by 5–20% (w/v) sucrose gradient centrifugation as described previously (Yamada et al., 1987), except that RNA samples were centrifuged at 24 000 rpm for 14 h at 4 °C on a Beckman SW 41T rotor. Double-stranded cDNA was synthesized by the method of Gubler and Hoffman (1983), and a cDNA library was constructed in λgt10 from the cDNA as described by Huynh et al. (1985). Briefly, the first DNA strand was synthesized from the enriched poly(A+) RNA (3 μg) with oligo(dT)$_{12-18}$ and reverse transcriptase in 20 μL of reaction mixture. The second strand was synthesized from the first strand with RNase H, DNA polymerase I, and *E. coli* DNA ligase. After treatment with T4 DNA polymerase, the cDNA was methylated with *Eco*RI methylase and then ligated with *Eco*RI linker. The cDNA was digested with *Eco*RI restriction endonuclease and fractionated by gel filteration on Bio-Gel A-50 m. The cDNA in the front half of the first peak was recovered by ethanol precipitation and ligated with *Eco*RI-digested λgt10 DNA. The ligated DNA was then packaged with Packergene and used to infect C600Hfl cells. The constructed cDNA library consisted of 1.7 × 10^6 independent recombinant phages.

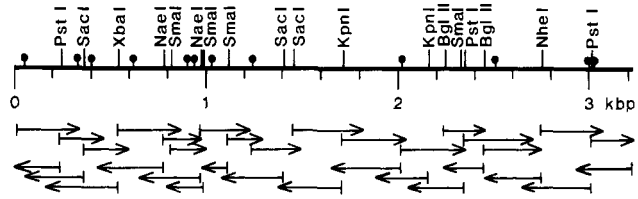


FIGURE 1: Restriction enzyme map of cDNA of λMP-H17 and sequencing strategy. The solid line shows the physical length of the cDNA in kilobase pairs. The first nucleotide begins at an initiation site of transcription. Closed circles on the line indicative *Pvu*II sites. Arrows parallel to the solid line indicate the direction and extent of nucleotide of sequencing.

The library was screened with a ^{32}P-labeled *Pst*I fragment (654 bp) of the cDNA from the myeloperoxidase cDNA clone pMP1 by plaque hybridization as described previously (Yamada et al., 1987). In all, 159 cDNA clones showing strong hybridization signals were selected, and several clones with cDNAs of more than 2.5 kbp were selected. The size of cDNA was determined by agarose gel electrophoresis. The *Pst*I fragment was labeled with [α-^{32}P]dCTP by the random oligoprimer method (Feinberg & Vogelstein, 1983).

Cloning of Myeloperoxidase Genomic DNA. A human genomic library in Charon 4A (Lawn et al., 1978) was a gift from Dr. T. Maniatis. About 1.8 × 10^6 phages were screened by the plaque hybridization described by Lawn et al. (1978) with a ^{32}P-labeled cDNA (3207 bp) and a ^{32}P-labeled *Eco*RI–*Xba*I fragment (530 bp) from the 5′ region of cDNA from λMP-H17 isolated as described above. Several genomic clones were selected. They were amplified and purified by the CsCl banding method (Maniatis et al., 1982).

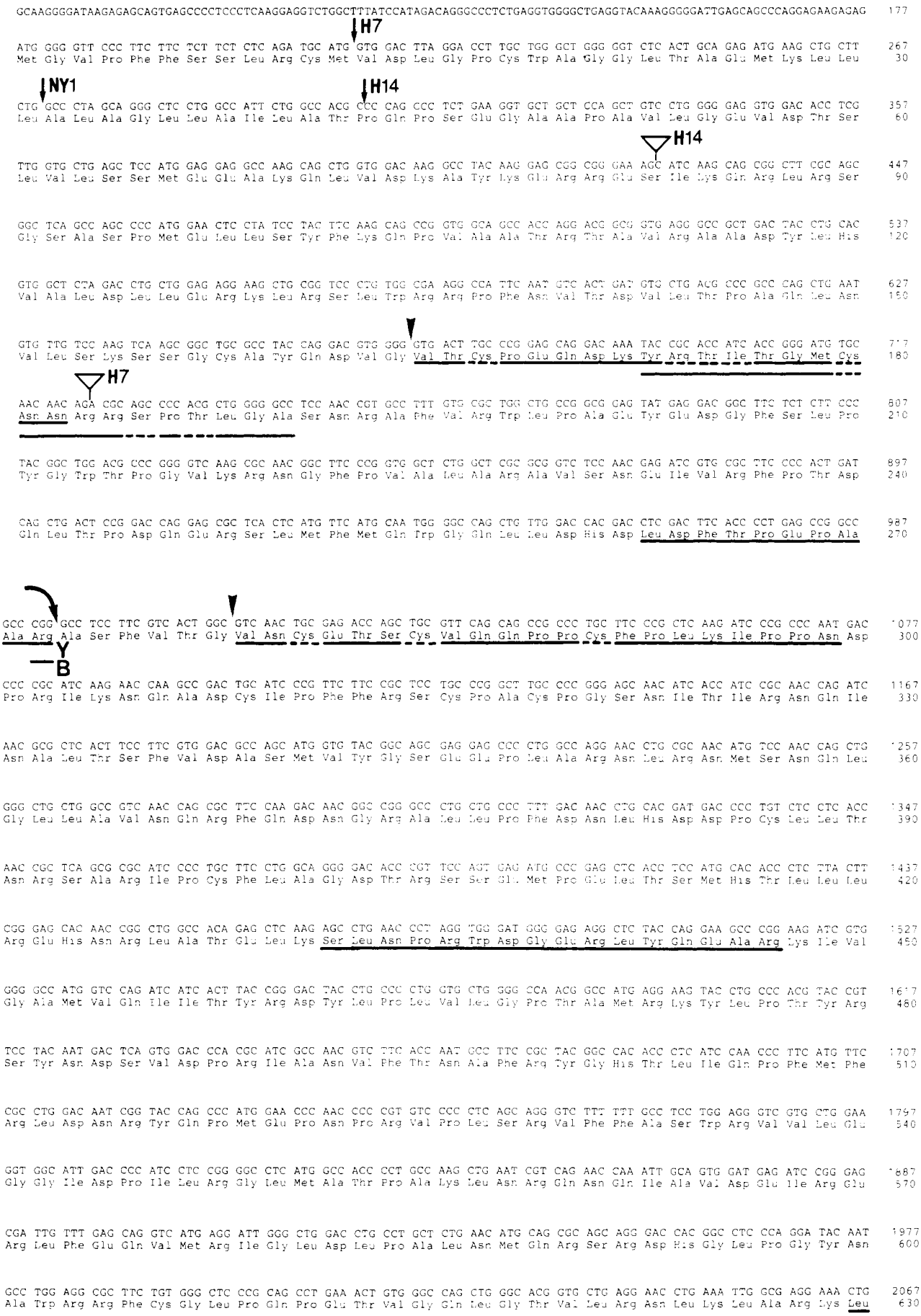
DNA Sequencing. The insert DNAs of cDNA clones and genomic clones were subcloned in pUC19. The insert DNA was digested with various restriction endonucleases, and then the restriction fragments were subcloned into M13 mp10, mp11, mp18, or mp19 and the DNAs were isolated from the recombinant clones by the method of Messing (1983). The DNA was sequenced with [α-^{35}S]dCTP by the dideoxy chain termination method (Sanger et al., 1977).

Primer Extension Analysis. Enriched poly(A+) RNA (1.8–6.9 μg) and a 17-base oligodeoxynucleotide (ATCCAGCTTCCAAGGAC, 3.3 pmol) were annealed in 2 μL of a solution of 125 mM Tris-HCl, pH 8.3, 25 mM $MgCl_2$, and 12.5 mM dithiothreitol at 60 °C for 1 h and then cooled to room temperature. The solution was mixed with a solution (3.8 μL) containing 8.8 μCi of [α-^{35}S]dCTP (13.5 pmol), 6.25 nmol each of dATP, dGTP, and dTTP, 40 nmol of sodium pyrophosphate, and 39 units of reverse transcriptase and incubated at 42 °C for 20 min. The reaction was chased with 0.12 μL of 25 mM dCTP for 15 min and then stopped by the addition of 4 μL of 96% formamide solution containing 20 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol.

The sample was boiled for 3 min, loaded on the top of an 8% polyacrylamide/8 M urea sequencing gel, and electrophoresed under the conditions used for sequencing. The nucleotide length of the primer extension products was determined with reference to a sequence ladder of the 5′ region of a genomic DNA formed by sequencing with the same primer as used for primer extension.

Amino Acid Sequencing of the Amino-Terminal and the Carboxy-Terminal Portions of the Light and Heavy Chains of Myeloperoxidase. Human leukocyte myeloperoxidase III

[1] Abbreviations: (k)bp, (kilo)base pair(s); SDS, sodium dodecyl sulfate; b, base; (k)Da, (kilo)dalton(s).



GCAAGGGGATAAGAGAGCAGTGAGCCCCTCCCTCAAGGAGGTCTGGCTTTATCCATAGACAGGGCCCTCTGAGGTGGGGCTGAGGTACAAAGGGGGATTGAGCAGCCCAGGAGAAGAGAG 177

ATG GGG GTT CCC TTC TTC TCT TCT CTC AGA TGC ATG GTG GAC TTA GGA CCT TGC TGG GCT GGG GGT CTC ACT GCA GAG ATG AAG CTG CTT 267
Met Gly Val Pro Phe Phe Ser Ser Leu Arg Cys Met Val Asp Leu Gly Pro Cys Trp Ala Gly Gly Leu Thr Ala Glu Met Lys Leu Leu 30

CTG GCC CTA GCA GGG CTC CTG GCC ATT CTG GCC ACG CCC CAG CCC TCT GAA GGT GCT GCT CCA GCT GTC CTG GGG GAG GTG GAC ACC TCG 357
Leu Ala Leu Ala Gly Leu Leu Ala Ile Leu Ala Thr Pro Gln Pro Ser Glu Gly Ala Ala Pro Ala Val Leu Gly Glu Val Asp Thr Ser 60

TTG GTG CTG AGC TCC ATG GAG GAG GCC AAG CAG CTG GTG GAC AAG GCC TAC AAG GAG CGG CGG GAA AGC ATC AAG CAG CGG CTT CGC AGC 447
Leu Val Leu Ser Ser Met Glu Glu Ala Lys Gln Leu Val Asp Lys Ala Tyr Lys Glu Arg Arg Glu Ser Ile Lys Gln Arg Leu Arg Ser 90

GGC TCA GCC AGC CCC ATG GAA CTC CTA TCC TAC TTC AAG CAG CCG GTG GCA GCC ACC AGG ACG GCG GTG AGG GCC GCT GAC TAC CTG CAC 537
Gly Ser Ala Ser Pro Met Glu Leu Leu Ser Tyr Phe Lys Gln Pro Val Ala Ala Thr Arg Thr Ala Val Arg Ala Ala Asp Tyr Leu His 120

GTG GCT CTA GAC CTG CTG GAG AGG AAG CTG CGG TCC CTG TGG CGA AGG CCA TTC AAT GTC ACT GAT GTG CTG ACG CCC GCC CAG CTG AAT 627
Val Ala Leu Asp Leu Leu Glu Arg Lys Leu Arg Ser Leu Trp Arg Arg Pro Phe Asn Val Thr Asp Val Leu Thr Pro Ala Gln Leu Asn 150

GTG TTG TCC AAG TCA AGC GGC TGC GCC TAC CAG GAC GTG GGG GTG ACT TGC CCG GAG CAG GAC AAA TAC CGC ACC ATC ACC GGG ATG TGC 717
Val Leu Ser Lys Ser Ser Gly Cys Ala Tyr Gln Asp Val Gly Val Thr Cys Pro Glu Gln Asp Lys Tyr Arg Thr Ile Thr Gly Met Cys 180

AAC AAC AGA CGC AGC CCC ACG CTG GGG GCC TCC AAC CGT GCC TTT GTG CGC TGG CTG CCG GCG CAG TAT GAG GAC GGC TTC TCT CTT CCC 807
Asn Asn Arg Arg Ser Pro Thr Leu Gly Ala Ser Asn Arg Ala Phe Val Arg Trp Leu Pro Ala Glu Tyr Glu Asp Gly Phe Ser Leu Pro 210

TAC GGC TGG ACG CCC GGG GTC AAG CGC AAC GGC TTC CCG GTG GCT CTG GCT CGC GCG GTC TCC AAC GAG ATC GTG CGC TTC CCC ACT GAT 897
Tyr Gly Trp Thr Pro Gly Val Lys Arg Asn Gly Phe Pro Val Ala Leu Ala Arg Ala Val Ser Asn Glu Ile Val Arg Phe Pro Thr Asp 240

CAG CTG ACT CCG GAC CAG GAG CGC TCA CTC ATG TTC ATG CAA TGG GGC CAG CTG TTG GAC CAC GAC CTC GAC TTC ACC CCT GAG CCG GCC 987
Gln Leu Thr Pro Asp Gln Glu Arg Ser Leu Met Phe Met Gln Trp Gly Gln Leu Leu Asp His Asp Leu Asp Phe Thr Pro Glu Pro Ala 270

GCC CGG GCC TCC TTC GTC ACT GGC GTC AAC TGC GAG ACC AGC TGC GTT CAG CAG CCG CCC TGC TTC CCG CTC AAG ATC CCG CCC AAT GAC 1077
Ala Arg Ala Ser Phe Val Thr Gly Val Asn Cys Glu Thr Ser Cys Val Gln Gln Pro Pro Cys Phe Pro Leu Lys Ile Pro Pro Asn Asp 300

CCC CGC ATC AAG AAC CAA GCC GAC TGC ATC CCG TTC TTC CGC TCC TGC CCG GCT TGC CCC GGG AGC AAC ATC ACC ATC CGC AAC CAG ATC 1167
Pro Arg Ile Lys Asn Gln Ala Asp Cys Ile Pro Phe Phe Arg Ser Cys Pro Ala Cys Pro Gly Ser Asn Ile Thr Ile Arg Asn Gln Ile 330

AAC GCG CTC ACT TCC TTC GTG GAC GCC AGC ATG GTG TAC GGC AGC GAG GAG CCC CTG GCC AGG AAC CTG CGC AAC ATG TCC AAC CAG CTG 1257
Asn Ala Leu Thr Ser Phe Val Asp Ala Ser Met Val Tyr Gly Ser Glu Glu Pro Leu Ala Arg Asn Leu Arg Asn Met Ser Asn Gln Leu 360

GGG CTG CTG GCC GTC AAC CAG CGC TTC CAA GAC AAC GGC CGG GCC CTG CTG CCC TTT GAC AAC CTG CAC GAT GAC CCC TGT CTC CTC ACC 1347
Gly Leu Leu Ala Val Asn Gln Arg Phe Gln Asp Asn Gly Arg Ala Leu Leu Pro Phe Asp Asn Leu His Asp Asp Pro Cys Leu Leu Thr 390

AAC CGC TCA GCG CGC ATC CCC TGC TTC CTG GCA GGG GAC ACC CGT TCC AGT GAG ATG CCC GAG CTC ACC TCC ATG CAC ACC CTC TTA CTT 1437
Asn Arg Ser Ala Arg Ile Pro Cys Phe Leu Ala Gly Asp Thr Arg Ser Ser Glu Met Pro Glu Leu Thr Ser Met His Thr Leu Leu Leu 420

CGG GAG CAC AAC CGG CTG GCC ACA GAG CTC AAG AGC CTG AAC CCT AGG TGG GAT GGG GAG AGG CTC TAC CAG GAA GCC CGG AAG ATC GTG 1527
Arg Glu His Asn Arg Leu Ala Thr Glu Leu Lys Ser Leu Asn Pro Arg Trp Asp Gly Glu Arg Leu Tyr Gln Glu Ala Arg Lys Ile Val 450

GGG GCC ATG GTC CAG ATC ATC ACT TAC CGG GAC TAC CTG CCC CTG GTG CTG GGG CCA ACG GCC ATG AGG AAG TAC CTG CCC ACG TAC CGT 1617
Gly Ala Met Val Gln Ile Ile Thr Tyr Arg Asp Tyr Leu Pro Leu Val Leu Gly Pro Thr Ala Met Arg Lys Tyr Leu Pro Thr Tyr Arg 480

TCC TAC AAT GAC TCA GTG GAC CCA CGC ATC GCC AAC GTC TTC ACC AAT GCC TTC CGC TAC GGC CAC ACC CTC ATC CAA CCC TTC ATG TTC 1707
Ser Tyr Asn Asp Ser Val Asp Pro Arg Ile Ala Asn Val Phe Thr Asn Ala Phe Arg Tyr Gly His Thr Leu Ile Gln Pro Phe Met Phe 510

CGC CTG GAC AAT CGG TAC CAG CCC ATG GAA CCC AAC CCC CGT GTC CCC CTC AGC AGG GTC TTT TTT GCC TCC TGG AGG GTC GTG CTG GAA 1797
Arg Leu Asp Asn Arg Tyr Gln Pro Met Glu Pro Asn Pro Arg Val Pro Leu Ser Arg Val Phe Phe Ala Ser Trp Arg Val Val Leu Glu 540

GGT GGC ATT GAC CCC ATC CTC CGG GGC CTC ATG GCC ACC CCT GCC AAG CTG AAT CGT CAG AAC CAA ATT GCA GTG GAT GAG ATC CGG GAG 1887
Gly Gly Ile Asp Pro Ile Leu Arg Gly Leu Met Ala Thr Pro Ala Lys Leu Asn Arg Gln Asn Gln Ile Ala Val Asp Glu Ile Arg Glu 570

CGA TTG TTT GAG CAG GTC ATG AGG ATT GGG CTG GAC CTG CCT GCT CTG AAC ATG CAG CGC AGC AGG GAC CAC GGC CTC CCA GGA TAC AAT 1977
Arg Leu Phe Glu Gln Val Met Arg Ile Gly Leu Asp Leu Pro Ala Leu Asn Met Gln Arg Ser Arg Asp His Gly Leu Pro Gly Tyr Asn 600

GCC TGG AGG CGC TTC TGT GGG CTC CCG CAG CCT GAA ACT GTG GGC CAG CTG GGC ACG GTG CTG AGG AAC CTG AAA TTG GCG AGG AAA CTG 2067
Ala Trp Arg Arg Phe Cys Gly Leu Pro Gln Pro Glu Thr Val Gly Gln Leu Gly Thr Val Leu Arg Asn Leu Lys Leu Ala Arg Lys Leu 630

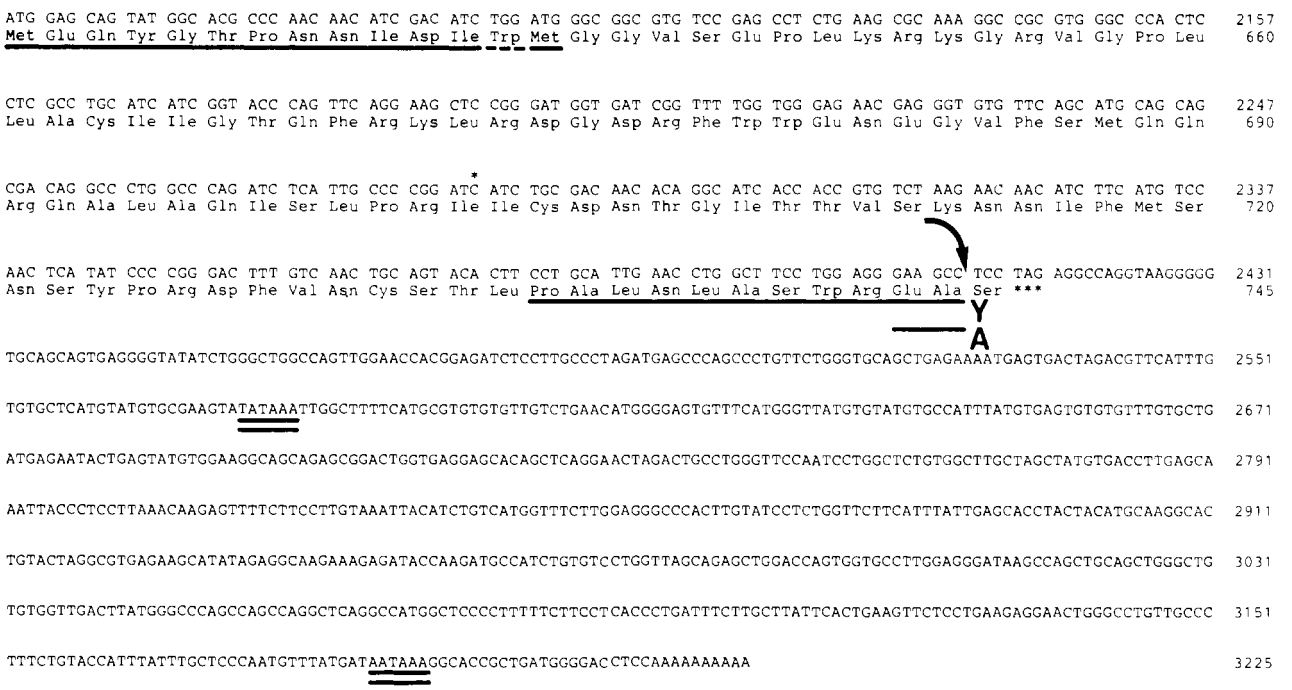


FIGURE 2: Nucleotide sequence of cDNA of λMP-H17 and deduced amino acid sequence of the myeloperoxidase. The nucleotide sequence in parentheses at the 5′ end was determined from the genomic DNA. Arrows H7, NY1, and H14 show the 5′ ends of the cDNAs of λMP-H7, λMP-NY1, and λMP-H14, respectively. Triangles H14 and H7 show the positions of the insertion of an extra sequence for the cDNAs of λMP-H14 and λMP-H7. Arrow heads show the amino termini of the light and heavy chains of myeloperoxidase; curved arrows show the carboxy termini of the light and heavy chains for the enzyme. Underlines show the amino acid sequences determined for the myeloperoxidase light and heavy chains and the lysylendopeptidase peptides from the light and heavy chains, and broken underlines show the amino acid residues unidentified; underlines Y, B, and A are the sequences determined with carboxypeptidase Y, B, and A, respectively; (***) shows the stop codon; the double underlines below nucleotide sequences indicate polyadenylation signal sequences; (*) nucleotide C was replaced by nucleotide T in cDNA of pMP1 (Yamada et al., 1987).

was purified as described previously (Suzuki et al., 1986). The two subunits were separated by SDS–polyacrylamide gel electrophoresis and eluted from the gel as described previously (Yamada et al., 1987). The amino-terminal amino acid sequence was determined in an Applied Biosystem 470A protein sequencer (Hirado et al., 1985). For determination of the carboxy-terminal amino acid sequence, the subunit (5 nmol) was digested with a molar ratio of carboxypeptidase Y, A, or B to the substrate of 1/100 to 1/500 in 0.1% SDS solution containing 10 mM sodium phosphate (pH 6.5) for carboxypeptidase Y or 0.2 M *N*-ethylmorpholine acetate (pH 8.2) for carboxypeptidase A and B at 37 °C, and aliquots containing 1 nmol were removed at intervals. The amino acids released were determined in a Hitachi 835S amino acid analyzer. The two subunits were also digested with lysylendopeptidase, the digests were separated by high-performance liquid chromatography (Yamada et al., 1987), and the amino acid sequences of the subunits were determined.

RESULTS

Cloning and Sequence Analysis of Various Full-Length cDNAs Encoding Myeloperoxidase. A cDNA library of HL-60 cells was constructed in λgt10. About 4.8×10^5 recombinant phages were screened by use of the cDNA encoding the carboxy terminus of myeloperoxidase as a probe, and 254 positive clones were isolated. Four of these that contained a large cDNA were selected for sequence analysis. The clone that had the largest cDNA insert is referred to as λMP-H17, and the three other clones are referred to as λMP-H7, λMP-NY1, and λMP-H14.

The cDNA of λMP-H17 was sequenced. Figure 1 shows the restriction endonuclease map and the sequencing strategy for the cDNA. Both strands were sequenced, and the results are shown in Figure 2. The sequence revealed that the cDNA was 3207 bp long and contained a 5′ noncoding regioin of 159 bp, and open reading frame of 2238 bp, and a 3′ noncoding region of 810 bp. The 5′ noncoding region contained the first ATG codon and consisted of a short open reading frame of 36 bp. The long open reading frame, starting at the second ATG codon of the cDNA and ending at a stop codon (TAG), coded for a 745 amino acid polypeptide. The amino acid sequence deduced from the nucleotide sequence is also shown in Figure 2. The 3′ noncoding region contained a polyadenylation signal, AATAAA, and ended with a poly(A) tail of 10 bp.

The amino acid sequences of the amino-terminal portions of the light and heavy chains of myeloperoxidase and of several peptides produced by lysylendopeptidase digestions of the light and the heavy chains were also determined. The carboxy-terminal sequence of the light chain was determined by digestions with carboxypeptidase Y and B, and that of the heavy chain was determined by digestions with carboxypeptidase Y and A. The results confirmed that the H17 cDNA encodes myeloperoxidase. Furthermore, it became evident that, for formation of the myeloperoxidase light chain (108 amino acid residues) and heavy chain (466 amino acid residues), a large amino-terminal polypeptide (164 amino acid residues), a small peptide (6 amino acid residues) between the light and the heavy chains, and a single amino acid at a carboxy terminus were removed from a primary polypeptide translated by the mRNA (Figure 2).

Structures of cDNAs with Alternative Exons. The four cDNA clones selected as described above were subcloned into the *Eco*RI site of pUC19 and characterized by comparison of the sizes of the fragments produced by digestions with various restriction enzymes. First, both ends of these cDNAs were sequenced. The sequences of the 5′ end regions were all included in the 5′ region of the λMP-H17 cDNA sequence.

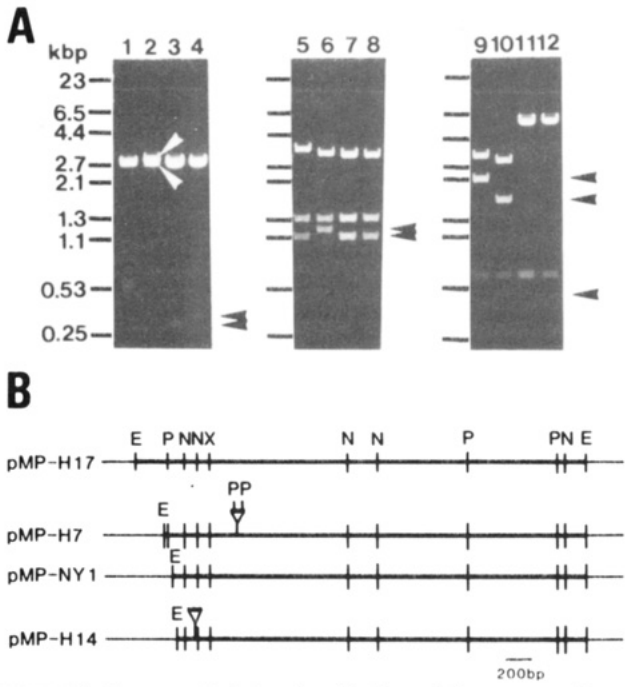


FIGURE 3: Agarose gel electrophoresis of restriction enzyme digests of various cDNAs. (A) cDNAs of four clones subcloned in the *Eco*RI site of pUC19 were digested with *Eco*RI and *Xba*I, *Nco*I, or *Pst*I. the digests were electrophoresed on 1.0% agarose gel and stained with ethidium bromide. (Lanes 1, 5, and 9) pMP-H17; (lanes 2, 6, and 10) pMP-H7; (lanes 3, 7, and 11) pMP-NY1; (lane 4, 8, and 12) pMP-H14. (Lanes 1–4) Digests with *Eco*RI and *Xba*I; (lanes 5–8) *Nco*I digests; (lanes 9–12) *Pst*I digests. The positions of λ–*Hin*dIII fragments and of pMP-H17 digests are shown with reference to a standard size marker in kilobase pairs. Arrowheads indicate bands with different mobilities. (B) Restriction enzyme maps of four cDNAs from plasmids H17, H7, NY1, and H14. The thick line shows cDNA and the thin line vector DNA. Triangles above lines show the insertion sites of extra sequences of H7 and H14 cDNAs determined from sequencing data. E, *Eco*RI; N, *Nco*I; P, *Pst*I; X, *Xba*I.

But their 5′ ends were located at different sites in H17 cDNA, as shown in Figures 2 and 3B. The 3′ end regions determined were all identical with that of the H17 cDNA with a slight difference of several nucleotides for the polyadenylation site. From the results, the sizes of the cDNAs were expected to be in the order H17 cDNA, H7 cDNA, NY1 cDNA, and H14 cDNA.

Agarose gel electrophoresis analysis of the sizes of these cDNAs excised by *Eco*RI digestion revealed that H14 cDNA was a little larger than NY1 cDNA (data not shown), suggesting that there is an extra sequence in the H14 cDNA. For determination of the presence or absence of an extra sequence in the other cDNAs, the DNAs of these plasmid clones were digested with various restriction enzymes specific for the cleavage of the cDNAs, and their digests were analyzed by agarose gel electrophoresis. The results are shown in Figure 3A. The restriction fragments of the four plasmid DNAs produced by digestions with *Eco*RI and *Xba*I are shown in lanes 1–4. The digest of plasmid H17 DNA gave two bands of about 2.7 kb and 530 b (lane 1). The 2.7-kb band included two fragments from the vector DNA and the cDNA. The plasmid H7 digest gave two strong bands of about 2.8 and 2.7 kb and a faint band of 340 b (lane 2). The 2.8-kb fragment was not seen in the plasmid H17 digest and could not be expected from the sequence of H17 cDNA (Figure 3B), suggesting the presence of an extra sequence in the fragment of H7 cDNA. The digests of pMP-NY1 and pMP-H14 each gave two bands of about 2.7 kb and 300 b (lanes 3 and 4). The small fragment from the NY1 digest was distinctly smaller than that from the H14 digest (lanes 3 and 4), suggesting the presence of an extra sequence in the fragment of H14 cDNA. The large fragments of the two cDNAs were the same as that of H17 cDNA.

The *Nco*I digests of the four plasmids each gave three strongly stained bands (lanes 5–8). The largest bands of 3.0–3.3 kb corresponded to a fragment containing a 5′ region of cDNAs and a vector DNA portion. The second-largest bands of the four digests were all the same size (lanes 5–8). The third band of the H7 digest of 1.2 kb was larger than that of the H17 digest (lanes 5 and 6), suggesting the presence of an extra sequence in the 1.2-kb fragment of H7 cDNA. The third bands of the NY1 and H14 digests were the same size as that of the H17 digest (lanes 5, 7, and 8).

The *Pst*I fragments of the digests of the four plasmids are shown in lanes 9–12. The 2.1-kb band was seen in the H17 digest (lane 9), but the H7 digest showed two bands of 1.6 kbp and 500 bp instead (lane 10), indicating that at least one additional *Pst*I site was present in the extra sequence of H7 cDNA. The 650-b band was seen in all four plasmid digests. These results revealed that H7 and H14 cDNAs had an extra sequence not present in H17 and NY1 cDNAs and that it was located in different sites (Figure 3B).

Next, the two extra sequences and their flanking regions were sequenced, and their exact locations were determined. For sequence determination of the H7 insert, plasmid H7 DNA was digested with *Pvu*II or with *Sac*I and *Nae*I, and then the *Pvu*II fragment of 376 bp and the *Sac*I–*Nae*I fragment of 503 bp were isolated from the digests and were subcloned into M13 DNA. For sequence determination of the H14 insert, plasmid H14 was digested with *Eco*RI and *Xba*I, and then the *Eco*RI–*Xba*I fragment of 327 bp was isolated from the digest and subcloned into M13 DNA. The sequencing strategies and the length of the sequences determined of H7 and H14 cDNAs are shown in Figure 6. The sequences determined are shown in Figure 4. The sequence from H7 cDNA was 96 bp and was located on the 726th nucleotide of H17 cDNA. The sequence from H14 cDNA was 82 bp and was located on the 426th nucleotide. The 5′ and 3′ flanking sequences of the H7 and H14 inserts were each identical with the sequences located on both sides of the 725th nucleotide and the 425th nucleotide, respectively, of H17 cDNA.

H7 cDNA contained the same reading frame as H17 cDNA, but it was longer and encoded a 777 amino acid polypeptide, whereas H17 cDNA encoded a 745 amino acid polypeptide. With the 82-b sequence insertion in H14 cDNA, the open reading frame started at the second ATG codon in the same way as in H17 cDNA, but was interrupted with a stop codon shortly after the start (Figure 4B). A long open reading frame could start at the sixth ATG codon and was identical with the frame of H17 cDNA. Therefore, the cDNA could encode a 650 amino acid polypeptide. The polypeptide still included a leader peptide (69 amino acid residues) and the myeloperoxidase light and heavy chains.

Structure of a Myeloperoxidase Genomic DNA. For clarification of the origins of the three cDNAs, myeloperoxidase genomic DNA was isolated from a human gene library and characterized. Two clones were selected. They both contained genomic DNAs of about 15 kb. They gave the same patterns of restriction fragments with various restriction enzymes, except that the 5′ flanking region of the gene was longer in one clone than in the other. The former, named λMP018, was sequenced. The sequence comprising the 5′ flanking region and the 5′ region of the genomic DNA was determined (Figure 5).

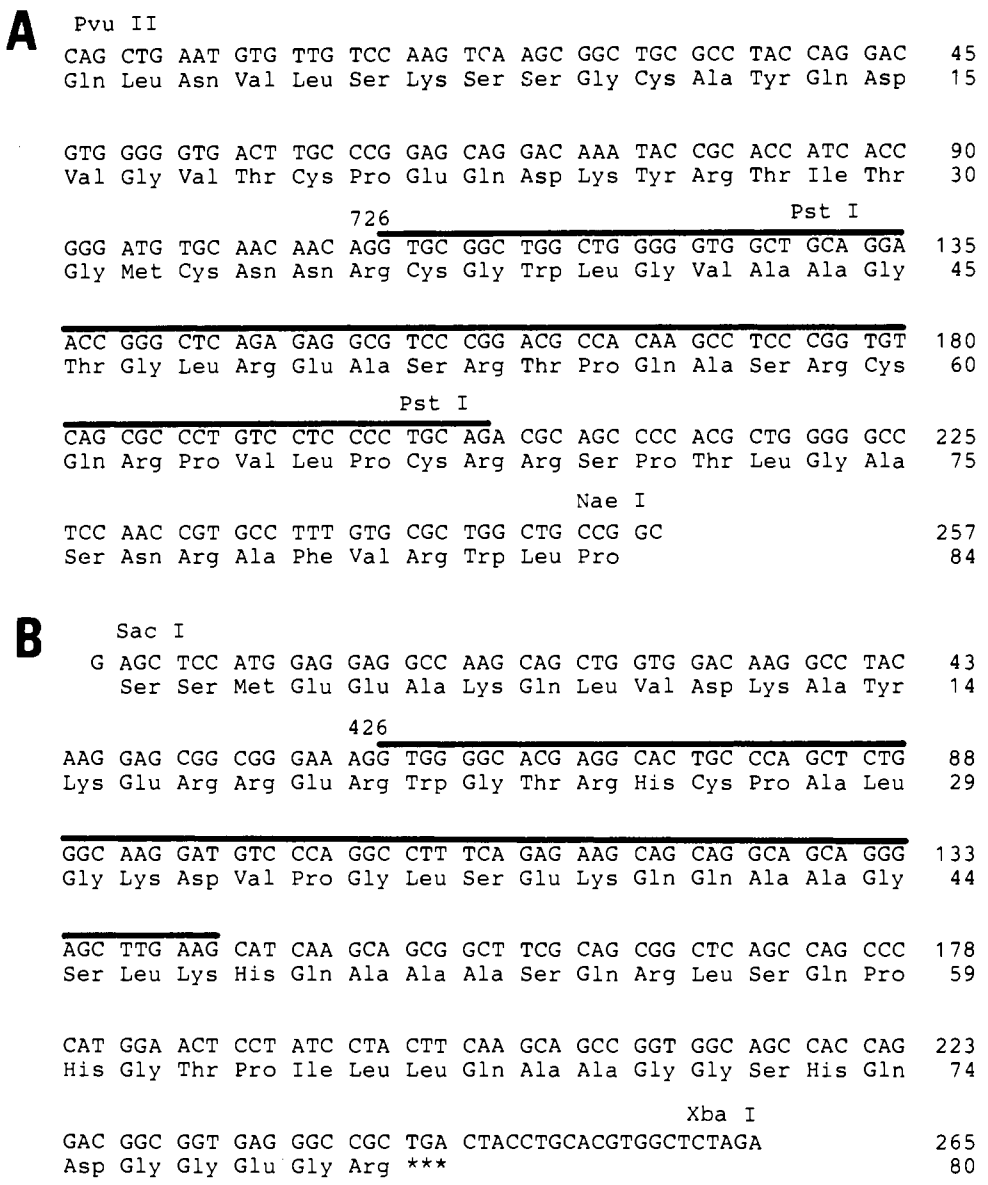


FIGURE 4: Nucleotide sequences of short inserts and their flanking region in H7 and H14 cDNAs and the deduced amino acid sequence. Lines above the nucleotide sequence indicate the inserted sequences in H7 cDNA (A) and H14 cDNA (B). The numbers 726 and 426 at the beginning of the lines indicate the positions of the first nucleotides of the H7 and H14 inserts, respectively, on H17 cDNA. (***) shows the stop codon.

Examination of the nucleotide sequences showed that those of the three cDNAs including the two extra sequences were all in the 5′ region of the genomic DNA. Furthermore, the 5′ region of the gene was split into four exons and three introns (Figure 6). Parts of exon 2 and 4 were optional in these three cDNAs. Exon 2 was composed of two contiguous segments. The former half of the exon was present in all three cDNAs, while the latter half was found in H14 cDNA but not in the other cDNAs. Exon 4 was composed of three contiguous segments. The first segment and the third segment were found in all three cDNAs, but the second one was found in only H7 cDNA. These results indicated that these cDNAs could be generated by alternative splicing from a single transcript of a single gene.

Structure of the 3′ Region of cDNAs with Different Polyadenylation Sites. Five classes of the 3′ ends of cDNA clones are shown in Table I. During sequencing of the 3′ ends of the large cDNA clones described above, two classes of polyadenylation site were found 15- and 22-bp, respectively, downstream from the polyadenylation signal AATAAA located at the 3188th nucleotide (Table I). Another polyadenylation signal, TATAAA, is known to be present 613-bp upstream from this signal (Morishita et al., 1987a). Therefore, for examination of this polyadenylation signal, several cDNA clones were digested with *Eco*RI. The digests were electro-

Table I: Multiple Polyadenylation Sites for Myeloperoxidase mRNA[a]

polyadenylation signal sequences and polyadenylation sites	clones
TYPE I	
---GCGAAGTA(2575)TATAAA(10 b)$(A)_{20}$	λMP-NY20
---GCGAAGTATATAAA(14 b)TGGCTTTTCATGC$(A)_{13}$	λMP-H123
---GCGAAGTATATAAA(17 b)TGGCTTTTCATGCGTG$(A)_{21}$	λMP-H80, λGM706[b]
TYPE II	
---TTTATGATA(3188)ATAAA(15 b)GGCACCGCTGATGGG$(A)_{25}$	λMP-NY1, pMP-1[c]
---TTTATGATAATAAA(22 b)GGCACCGCTGATGGGGACCTCC$(A)_{10-17}$	λMP-H7, H10, H14, H17

[a] Classification of the nucleotide sequences of the 3′ regions of various cDNAs encoding myeloperoxidase. [b] Reported by Morishita et al. (1987a). [c] Reported by Yamada et al. (1987).

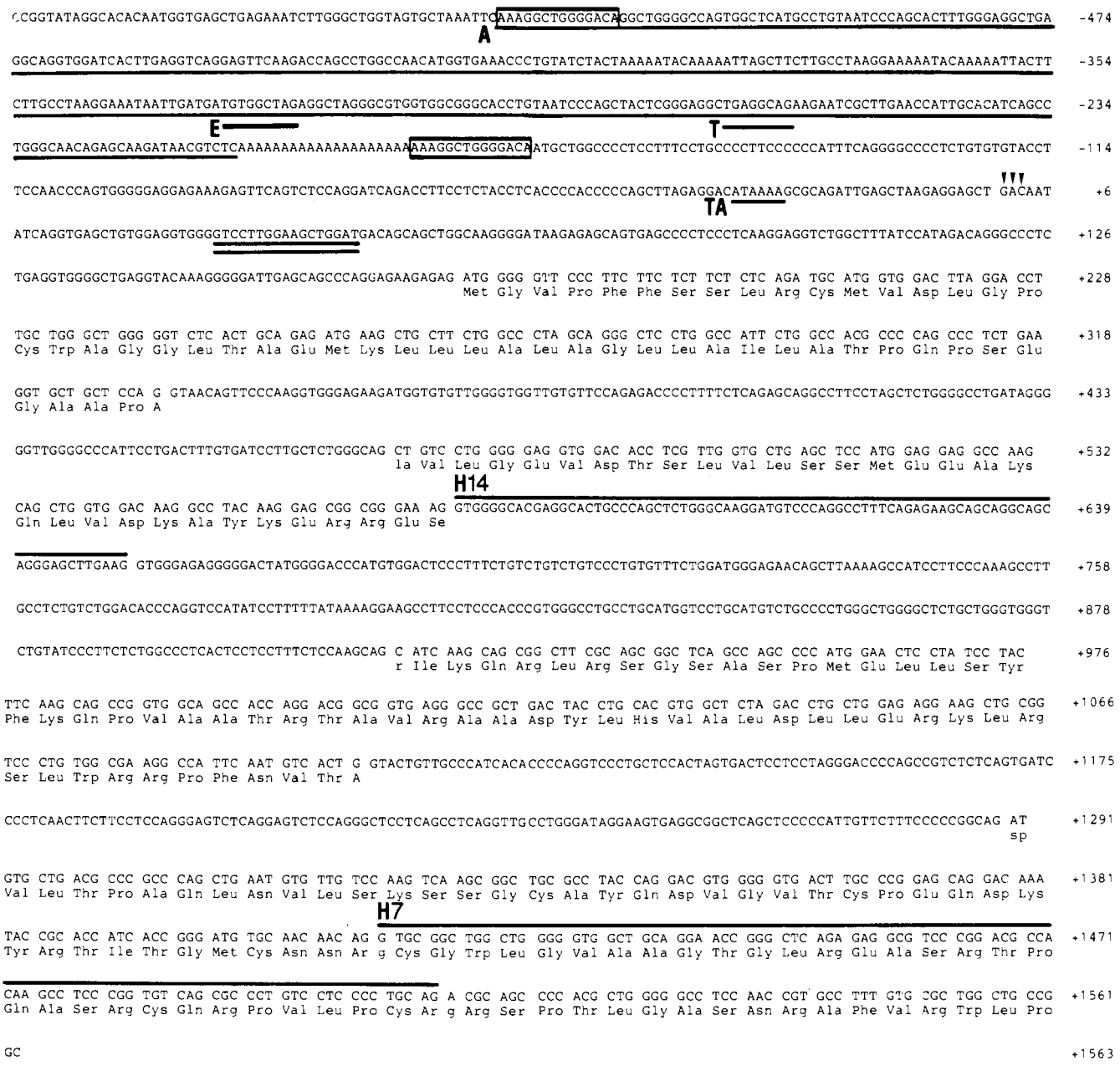


FIGURE 5: Sequence of the 5′ flanking region and 5′ region of myeloperoxidase genomic DNA. The genomic DNA of λMP018 was digested with various restriction enzymes, and the fragments were subcloned in M13 and sequenced. Underlines A, E, T, and TA show the *Alu* sequence, enhancer core sequence, phorbol ester responsive element-like sequence, and degenerated TATA box, respectively; boxed sequences indicate direct repeats flanking the *Alu* family; the double underline shows the sequence for the synthetic oligodeoxynucleotide used in primer extension analysis. Arrowheads show major positions of the 5′ ends of myeloperoxidase mRNA detected by primer extension analyses as described under Experimental Procedure. Lines H14 and H17 above the nucleotide sequence indicate short inserted sequences in H14 and H17 cDNAs, respectively.

phoresed on agarose gel and then hybridized to a ^{32}P-labeled *Eco*RI–*Xba*I fragment derived from the 5′ region of H17 cDNA. The positive clones with inserts of less than 2.7 kbp were selected, and their 3′ ends were sequenced. Three more classes of polyadenylation sites were found 10-, 14-, and 17-bp, respectively, downstream from the polyadenylation signal (Table I).

Discussion

In this work we isolated and characterized three types of full-length cDNAs encoding myeloperoxidase. Sequence analyses of these cDNAs indicated that λMP-H17 cDNA encodes a myeloperoxidase polypeptide of 745 amino acids residues. This encoded polypeptide is the same as those encoded by the cDNAs reported by others (Morishita et al., 1987a; Johnson et al., 1987). The other cDNA clone, λMP-H7, contained a small coding exon in the open reading frame of H17 cDNA and so encoded a polypeptide containing an additional 32 amino acid peptide in the middle of the myeloperoxidase light chain. At least three types of myeloperoxidase from human leukocytes and HL-60 cells are known. Myeloperoxidase III, among them, is a major type in HL-60 cells (Yamada et al., 1981). But as their primary structures are still not known, it is impossible to correlate them with the proteins predicted from the cDNAs. The amino acid sequence of a lysylendopeptidase peptide from the light chain of myeloperoxidase III suggests that myeloperoxidase III might be the protein encoded by H17 and H14 cDNAs, because in the H7 cDNA encoded protein the peptide should be interrupted at the arginine residue by addition of a unique sequence (Figure 2).

The sequence of H14 cDNA revealed that the open reading frame starts at the first ATG codon in exon 3 and can encode

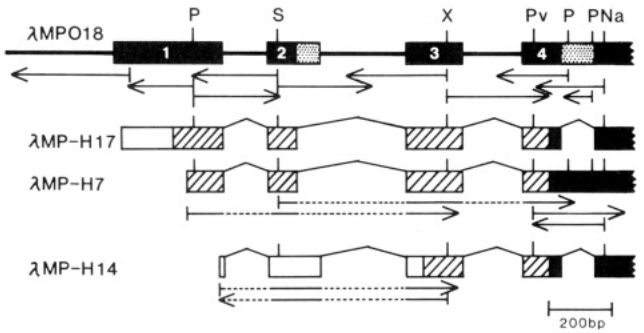


FIGURE 6: Structural organization of the myeloperoxidase gene 5′ region and the three spliced mRNAs. Black boxes of the gene from a genomic clone λMP018 show constitutive exons; dotted boxes show alternative sequences; solid lines show the 5′ flanking region and introns. Three patterns of alternative splicing of mRNA are shown parallel to the gene organization. The splicing pathway is trailed by connecting boxes of exon sequences by a thin line. Striped boxes indicate the part encoding the leader peptide of a precursor; black boxes show the part encoding the light chain, and white boxes show the 5′ noncoding exon. Splicing was shown by cDNA sequencing. Arrows indicate the direction of sequencing, and arrow lengths correspond to the extents of nucleotide sequences determined in each sequencing. Na, *Nae*I; P, *Pst*I; Pv, *Pvu*II; S, *Sac*I; X, *Xba*I. The *Pvu*II site on exon 4 only is shown for clarification.

a 650 amino acid polypeptide including the myeloperoxidase light and heavy chains. This polypeptide was the same as that encoded by H17 cDNA, except that the preceding peptide of the former began at the 96th amino acid residue on that of the latter. Therefore, the primary translation products from the mRNAs corresponding to these cDNAs should differ in size. However, when poly(A+) RNA from HL-60 cells was translated in a reticulocyte lysate, a myeloperoxidase polypeptide of approximately 74 kDa was detected as a single band on SDS-polyacrylamide gel electrophoresis (Yamada & Kurahashi, 1984). The first ATG codons of the open reading frames from H17 and H7 cDNAs were present in exon 1, and that from H14 cDNA was in exon 3. The latter ATG codon seems more plausible for initiation than the former because the sequence surrounding the initiation codon is the more homologous with the consensus sequence for initiation in higher eukaryotes found by Kozak (1987). But it is still unknown whether both these two putative initiator codons or only the initiator codon in exon 3 functions in translation of myeloperoxidase mRNA in protein synthesis.

Sequence analysis of a myeloperoxidase genomic DNA showed that these cDNAs all came from the genomic DNA. The gene organization of myeloperoxidase was reported by Morishita et al. (1987b) while this work was in progress. The structure of the 5′ region was the same as that found in this work. Futhermore, recently a single myeloperoxidase gene was found to be present in the human genome and to be located on chromosome 17 (Weil et al., 1987; Kudoh et al., 1988). Southern blot analysis of the *Eco*RI digest of human genomic DNA with myeloperoxidase cDNA as a probe showed the presence of a single *Eco*RI fragment of 20 kb (Chang et al., 1986; Morishita et al., 1987a; Kudoh et al., 1988). The present results taken together indicate that these cDNAs are all derived from a single gene by alternative use of the exons. The three patterns of myeloperoxidase RNA splicing shown in this work were also found in various patterns of alternative RNA splicing in a wide variety of mRNAs from the other sources (Breitbart et al., 1987).

Primer extension experiments indicated that the major 5′ end of myeloperoxidase mRNA is located 18 nucleotides upstream from the 5′ end of H17 cDNA (data not shown). A similar site for the 5′ end of the mRNA was also detected by S1 nuclease protection experiments (Morishita et al., 1987b). Thus, myeloperoxidase gene transcription maybe started mainly from the site. Examination of the 5′ flanking region for up to about 500 nucleotides upstream from the initiation site showed no authentic sequences of CAAT and TATA responsible for initiation of transcription by eukaryotic RNA polymerase II (Breathnach & Chambon, 1981). However, an *Alu* sequence (Schmid & Jelinek, 1982), an enhancer core sequence (Weiher et al., 1983), and a TPA responsive element-like sequence (Angel et al., 1987) were found. Studies are required on whether some of these sequences are associated with downregulation of myeloperoxidase gene expression in HL-60 cells by 12-*O*-tetradecanoylphorbol 13-acetate or retinoic acid (Kasugai & Yamada, 1986; Koeffler et al., 1985; Yamada & Kurahashi, 1984). Several classes of multiple initiation sites for transcription of the mouse band 3 gene and several other genes were observed, and their promoter regions also lacked CAAT and TATA and consisted of multiple GC-rich sequences (Kopito et al., 1987; Maire et al., 1987).

Two species of myeloperoxidase mRNA from HL-60 cells detected by Northern blot analysis were approximately 3.3 or 2.6 kb in length (Chang et al., 1986; Johnson et al., 1987; Weil et al., 1987). This size difference between the mRNAs could be interpreted mainly as due to alternative use of two polyadenylation signals, which were separated by approximately 600 nucleotides.

The present studies indicated that heterogeneous mRNAs coding for myeloperoxidase are formed in various ways from a single gene. The amounts of individual species of these mRNAs were not estimated in this work. Therefore, it is hard to know how much these heterogeneous mRNAs are related with the syntheses of multiple forms of myeloperoxidase. Further studies on the chemical structures of myeloperoxidases are required for clarification of this.

ACKNOWLEDGMENTS

We thank Dr. Tom Maniatis, Harvard University, for providing a human genomic library, and Dr. Earl R. Davie, University of Washington, for valuable advice. We also thank Taeko Kobayashi for typing the manuscript.

REFERENCES

Andrews, P. C., & Krinsky, N. I. (1981) *J. Biol. Chem. 256*, 4211–4218.

Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P., & Karin, M. (1987) *Cell (Cambridge Mass.) 49*, 729–739.

Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.

Breitbart, R. E., Andreadis, A., & Nadal-Ginard, B. (1987) *Annu. Rev. Biochem. 56*, 467–495.

Chang, K. S., Trujillo, J. M., Cook, R. G., & Stass, S. A. (1986) *Blood 68*, 1411–1414.

Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.

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

Henderson, W. R., & Klebanoff, S. J. (1983) *J. Biol. Chem. 258*, 13522–13527.

Hirado, M., Tsunasawa, S., Sakiyama, F., Niinobe, M., & Fujii, S. (1985) *FEBS Lett. 186*, 41–45.

Huynh, T. V., Young, R. A., & Davis, R. W. (1985) in *DNA Cloning* (Glover D. M., Ed.) Vol. 1, pp 49–78, IRL, Oxford.

Johnson, K. R., Nauseef, W. M., Care, A., Wheelock, M. J., Shane, S., Hudson, S., Koeffler, H. P., Selsted, M., Miller, C., & Rovera, G. (1987) *Nucleic Acids Res. 15*, 2013–2028.

Kasugai, I., & Yamada, M. (1986) *J. Biochem. (Tokyo) 100*, 381–388.
Klebanoff, S. J. (1975) *Semin. Hematol. 12*, 117–142.
Koeffler, H. P., Ranyard, J., & Pertcheck, M. (1985) *Blood 65*, 484–491.
Kopito, R. R., Andersson, M. A., & Lodish, H. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7149–7153.
Kozak, M. (1987) *J. Mol. Biol. 196*, 947–950.
Kudoh, J., Minoshima, S., Hashinaka, K., Nishio, C., Yamada, M., Shimizu, Y., & Shimizu, N. (1988) (*HGM9*) *Cytogenet. Cell Genet.* (in press).
Lawn, R. M., Fritsh, E. F., Parker, R. C., Blake, G., & Maniatis, T. (1978) *Cell (Cambridge Mass.) 15*, 1157–1174.
Maire, P., Gauton, S., Hakim, V., Gregori, C., Mennecier, F., & Kahn, A. (1987) *J. Mol. Biol. 197*, 425–438.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Messing, J. (1983) *Methods Enzymol. 101*, 20–78.
Miyasaki, K. T., Wilson, M. E., Cohen, E., Jones, P. C., & Genco, R. J. (1986) *Arch. Biochem. Biophys. 246*, 751–764.
Morishita, K., Kubota, N., Asano, S., Kaziro, Y., & Nagata, S (1987a) *J. Biol. Chem. 262*, 3844–3851.
Morishita, K., Tsuchiya, M., Asano, S., Kaziro, Y., & Nagata, S. (1987b) *J. Biol. Chem. 262*, 15208–15213.
Morita, Y., Iwamoto, H., Aibara, S., Kobayashi, T., & Hasegawa, E. (1986) *J. Biochem. (Tokyo) 99*, 761–770.
Pember, S. O., Shapira, R., & Kinkade, J. M., Jr. (1983) *Arch. Biochem. Biophys. 221*, 391–403.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schmid, C. W., & Jelinek, W. R. (1982) *Science (Washington D.C) 216*, 1065–1070.
Suzuki, K., Yamada, M., Akashi, K., & Fujikura, T. (1986) *Arch. Biochem. Biophys. 245*, 167–173.
Weiher, H., König, M., & Gruss, P. (1983) *Science (Washington D.C.) 219*, 626–631.
Weil, S. C., Rosner, G. L., Reid, M. S., Chisholm, R. L., Farber, N. M., Spitznagel, J. K., & Swanson, M. S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2057–2061.
Yamada, M., & Kurahashi, K. (1984) *J. Biol. Chem. 259*, 3021–3025.
Yamada, M., Mori, M., & Sugimura, T. (1981) *Biochemistry 20*, 766–771.
Yamada, M., Hur, S.-J., Hashinaka, K., Tsuneoka, K., Saeki, T., Nishio, C., Sakiyama, F., & Tsunasawa, S. (1987) *Arch. Biochem. Biophys. 255*, 147–155.

Interaction of the Heavy Chain of Scallop Myosin Heads with Skeletal F-Actin†

Jean-Pierre Labbé, Patrick Chaussepied, Jean Derancourt, and Ridha Kassab*
Centre de Recherche de Biochimie Macromoléculaire du CNRS, INSERM U.249, Université de Montpellier I, Route de Mende, BP 5051, 34033 Montpellier Cedex, France

Received November 17, 1987; Revised Manuscript Received April 1, 1988

ABSTRACT: The Ca^{2+}-regulated molluscan myosin from scallop muscle and its subfragment 1 (+LC) were cross-linked to skeletal and scallop F-actins by using 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide (EDC) under conditions avoiding protein aggregation. The elevated Mg^{2+}-ATPase of the covalent acto-S-1 was as high in the absence as in the presence of Ca^{2+}. In contrast, the cross-linking of the intact myosin to actin led to an enhancement of the Mg^{2+}-ATPase that was 2-fold higher in the presence than in the absence of Ca^{2+}. Calcium had no effect on the stimulated ATPase of the covalent complex formed with desensitized myosin. Cross-linked adducts between actin and the 31K or 24K heavy chain peptides present in the tryptic (63K–31K)-S-1 and (70K–24K)-S-1 were produced but not between actin and the corresponding N-terminal segments. Joining the former derivative to actin led to an inactive covalent complex, indicating that an irreversible Mg^{2+}-ATPase loss follows the production of this split S-1 [Szentkiralyi, E. M. (1987) *J. Muscle Res. Cell Motil. 8*, 349–357]. We further show that this ATPase inhibition is specifically associated with the heavy chain scission at the 63K–31K site as proteolysis within the 31K–24K difference segment by *Staphylococcus aureus* protease did not impair the S-1 enzymatic activities. The 17-residue N-terminal sequence we determined for the isolated 31K peptide was not shared by myosins from vertebrate striated muscles. It would reflect the observed unique proteolytic behavior of the scallop S-1 heavy chain and a particular structural relationship between the 63K–31K junction and the ATPase site. Finally, the isolated, renatured, soluble 24K fragment failed to interact with actin, in contrast with the homologous peptide from skeletal S-1. Conversly, actin did not bind to a new trypsin S-1 derivative formed in the presence of Mg^{2+}-ATP and having an intact N-terminal 68K segment but a degraded 24K moiety. These data imply that critical interactions between the associated NH_2- and COOH-terminal regions of the S-1 heavy chain are required for the specific attachment of actin to the scallop myosin head.

The contraction of molluscan striated muscles depends on the force-generating interactions between the actomyosin complex and Mg^{2+}-ATP in the presence of calcium. The specific binding of Ca^{2+} ions to a regulatory subunit of myosin (Szent-Gyorgyi et al., 1973; Chantler & Szent-Gyorgyi, 1980) induces a structural rearrangement of a vicinal essential light chain (Hardwicke & Szent-Gyorgyi, 1985) that alters the ATPase site conformation and results in the activation of the actomyosin Mg^{2+}-ATPase through acceleration by actin of the

† This research was supported by grants from the Centre National de la Recherche Scientifique and the Institut National de la Santé et de la Recherche Médicale (C.R.E. 5-11850).
* To whom correspondence should be addressed.

0006-2960/88/0427-5914$01.50/0