membranes of large vessels nor neighboring ectodermal or endodermal basement membranes were affected. It is possible that loss of basement membrane scaffolding to which anchorage-dependent endothelial cells are normally attached may be involved in the associated capillary involution.

Our findings open the possibility of the chemical manipulation of angiogenesis. For example, this new class of steroids, when administered with an appropriate heparin fragment, may be of potential therapeutic benefit in some diseases dominated by a pathologic process which is angiogenesis-dependent.

References and Notes

1. J. Folkman, R. Langer, R. J. Linhardt, C. Haudenschild, S. Taylor, *Science* **221**, 719 (1983).
2. N. J. Holbrook, *et al.*, *Neurosci. Res.* **6**, 75 (1981); S. Szabo, *Acta Endocrinol. (Copenhagen) Suppl.* **230**, 1 (1979).
3. R. Crum and J. Folkman, *J. Cell Biol.* **99**, 158a (1984).
4. Anti-angiogenic activity of heparin (in the presence of steroids) varies with the manufacturer and the heparin lot. For example, at a concentration 25 μg/10 μl, the percentage of embryos with avascular zones was: Sigma 25 percent; Hepar 53 percent; Abbott Panheprin 89 percent (in the presence of hydrocortisone, 50 μg/10 μl); ($P < 0.05$ for three experiments). Therefore, when used locally in the chick embryo or in the rabbit cornea, heparins from different sources can be made effective by simply increasing the concentration of a less active heparin. This is not always possible when heparin is administered orally to achieve systemic anti-angiogenesis. When orally administered heparin was used (with corticosteroids) against mouse tumors (1), very high doses of heparin increased tumor growth, presumably due to enhancement of angiogenesis. Only Panheprin (Abbott) and, for some tumors, heparin from Hepar, Inc. were anti-angiogenic at a sufficiently low dose to inhibit tumor growth. Panheprin is no longer being manufactured.
5. G. W. Liddle and K. L. Melmon, in *Textbook of Endocrinology*, R. H. Williams, Ed. (Saunders, Philadelphia, ed. 5, 1974), p. 244.
6. J. J. Castellot, Jr., L. V. Favreau, M. J. Karnovsky, R. D. Rosenberg, *J. Biol. Chem.* **257**, 11256 (1982).
7. J. Choay, *Seminars in Thrombosis and Hemostasis* **11** (part 2), 81 (1985).
8. D. E. Ingber, J. A. Madri, J. Folkman, *J. Cell Biol.* **101**, 105a (1985).
9. R. Auerbach, L. Kubai, D. Knighton, J. Folkman, *Dev. Biol.* **41**, 391 (1974).

10. We thank D. Ingber, W. Risau and B. Zetter for helpful discussions; N. Chen for advice on steroid structures and for model building; D. Lund for the corneal implantations; W. Wierenga (Upjohn, Kalamazoo, Mich.), for the steroids; J. Choay (Choay Institute, Paris) for the synthetic pentasaccharide and the chemically cleaved heparin hexasaccharide. We also thank F. Barker for suggesting the term angiostatic; G. Jackson for the skill and patience required to prepare more than 16,000 shell-less chick embryos; B. Sabga for help in preparation and implantation of the methylcellulose disks; and P. Breen for manuscript preparation. Supported by USPHS grant R01-CA37395 and by a grant to Harvard University from the Monsanto Co.

23 July 1985; accepted 21 October 1985

Promoter Region of the Human Harvey *ras* Proto-oncogene: Similarity to the EGF Receptor Proto-oncogene Promoter

Abstract. *Regulation of transcription of members of the* ras *gene family undoubtably plays an important role in controlling cellular growth. Examination of this level of regulation requires identification of the promoter regions of the* ras *proto-oncogenes. Four major transcriptional start sites were detected in the human Harvey* ras *1 proto-oncogene. The promoter region contains neither a TATA box nor a CAAT box in their characteristic upstream positions, has an extremely high G+C content (80 percent), and contains multiple GC boxes including seven CCGCCC repeats and three repeats of the inverted complement, GGGCGG. This region has strong promoter activity when placed upstream from the chloramphenicol acetyl transferase gene and transfected into monkey CV1 cells. In these ways the Harvey* ras *1 proto-oncogene promoter resembles the promoter of the gene encoding the epidermal growth factor (EGF) receptor. The similarity between the two proto-oncogene promoters may be relevant to the mechanism by which the expression of such "growth control" genes is regulated.*

SHUNSUKE ISHII
GLENN T. MERLINO
IRA PASTAN
Laboratory of Molecular Biology, Division of Cancer, Biology and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

The *ras* genes were first characterized as the transforming genes of the Harvey and Kirsten murine sarcoma viruses (v-Ha-*ras* and v-Ki-*ras*, respectively) (*1*). The H-*ras*, K-*ras*, N-*ras*, and *rho* genes constitute a family of conserved mammalian genes that encode structurally related proteins of approximately 21,000 daltons (p21) (*1–3*). Mutant *ras* genes encoding altered proteins are found in many human and rodent tumor cells and are capable of the malignant transformation of NIH 3T3 cells, an established murine cell line (*2, 4–9*). The mammalian *ras* gene products are associated with the cytoplasmic face of the cell membrane (*10, 11*), bind guanosine triphosphate (*12*), and exhibit a low level of guanosine triphosphatase activity (*13*). Little is known about the regulation of the expression of the *ras* genes. This is partly because the promoters of the *ras* genes have not been definitively identified even though the sequence at the 5′ terminus of some of the *ras* genes is known (*14, 15*).

We recently identified the promoter region of the gene coding for the proto-oncogene of the human epidermal growth factor (EGF) receptor (*16*). The EGF receptor is thought to be the cellular homolog of the v-*erbB* oncogene (*17–20*). The promoter region of the human EGF receptor proto-oncogene lacks a characteristic TATA box and CAAT box. It contains the repeat sequence CCGCCC, which is also found in the 21–base pair (bp) repeats of the simian virus (SV40) early promoter. We speculated that these features may be relevant to the mechanism by which the expression of some "growth control" genes is regulated (*16*). To investigate this possibility, we have identified and characterized the promoter region of another such gene—the human c-Ha-*ras* 1 gene.

To define the 5′ end of the c-Ha-*ras* 1 messenger RNA (mRNA), we first used primer extension analysis. A ^{32}P-labeled 51-bp Pvu II–Alu I DNA fragment (Fig. 1, nucleotides 1735 to 1684), which encodes a portion of the amino terminus of the *ras* protein, was hybridized to poly(A)$^+$ RNA of either the human epidermoid carcinoma cell line A431 or the mouse NIH 3T3 fibroblast cell line transformed by pEJ6.6 DNA, which contains the activated c-Ha-*ras* 1 gene of the human EJ bladder carcinoma (*21*). The primer was extended with reverse transcriptase, and the size of the products was determined by denaturing gel electrophoresis and autoradiography. When A431 mRNA was used, the most abundant extended products had lengths of 135, 142, 147, and 159 nucleotides (Fig. 2, lane 1). The mRNA of pEJ6.6-transformed NIH 3T3 cells served as template to synthesize the same set of extended products (Fig. 2, lane 2). The DNA sequence of these extended primers was determined and compared with the sequence of the c-Ha-*ras* 1 genomic clone reported earlier (*14, 15*). The comparison of the two sequences showed that an intron is located between nucleotides 577 and 1616 in the genomic clone (Fig. 1). The sequence at the exon 1–intron 1 boundary (nucleotide 576) is CGGT, whereas the intron 1–exon 2 boundary (nucleotide 1616) is AGGT. The sequences of the minor species of extended primers, which are longer than the four major species (see Fig. 2), are different from the sequence of the c-Ha-*ras* 1 genomic clone. We conclude that the

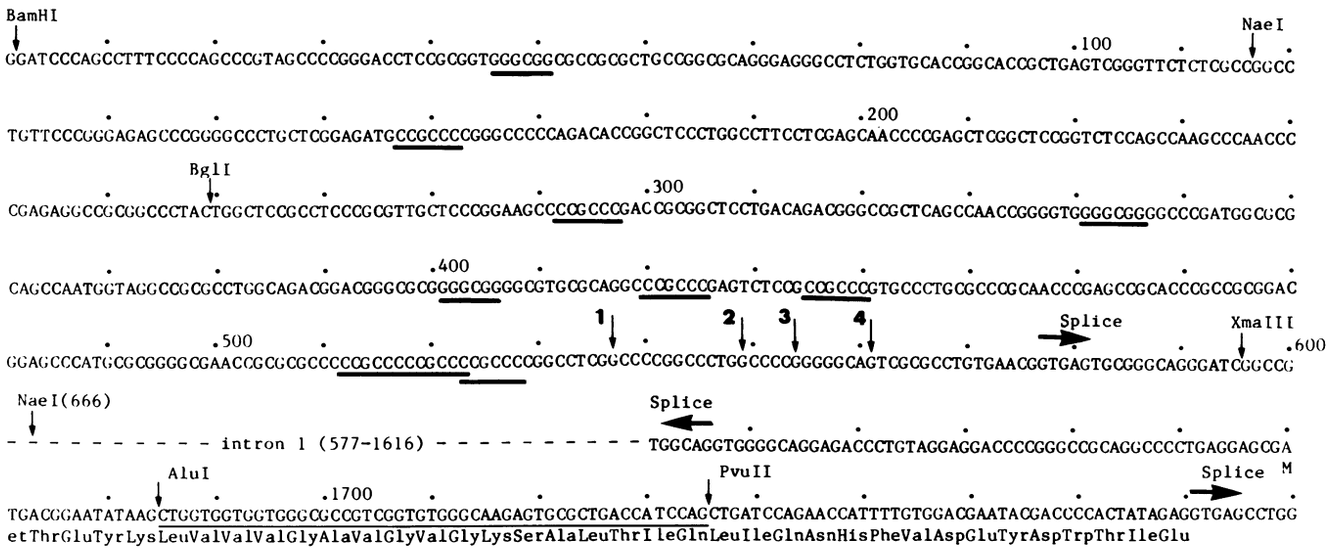


Fig. 1. Nucleotide sequence of the 5′ end of the human c-Ha-*ras* 1 gene (*14, 15*). The transcriptional initiation sites, determined by primer extension, are indicated by arrows 1 to 4. The Pvu II–Alu I DNA fragment (nucleotides 1735 to 1684) used as primer is underlined by a thin solid line. The repeated sequence CCGCCC and its complementary sequence GGGCGG are underlined by thick solid lines. The donor and acceptor splice signals are indicated.

minor species of extended primers are not generated from c-Ha-*ras* 1 mRNA. Figure 1 shows that the primers were extended to positions 537, 549, 554, and 561. These data are consistent with the likelihood that transcription of RNA initiates at four major sites; however, we cannot exclude the possibility that some of these result from reverse transcriptase strong stops. S1 nuclease mapping could not be performed at this region because of its proximity to the first intron (Fig. 1). The size of exon 1 is 40 bp (nucleotides 537 to 576) and it specifies only a 5′ untranslated segment of c-Ha-*ras* 1 mRNA. The 1040-bp first intron (nucleotides 577 to 1616) includes the region originally proposed to contain the promoter, as judged from sequence similarities to the canonical TATA and CAAT boxes (*14, 15*). The second exon is 164 bp long (nucleotides 1617 to 1780) and specifies 53 bp of the 5′ untranslated region and 111-bp coding for the first 37 amino acids of the p21 *ras* gene product.

To establish that the primer-extended region contains exon 1 and is transcribed in vivo, we fractionated A431 poly(A)$^+$ RNA electrophoretically, transferred the fractionated material to nitrocellulose, and hybridized it to either a ^{32}P-labeled Xma III–Bgl I DNA fragment (nucleotides 595 to 259) containing exon 1 sequences or to a ^{32}P-labeled Bgl I–Bam HI DNA fragment (nucleotides 259 to 2) containing sequences upstream from the observed transcriptional initiation region. Figure 3 shows that the 1.2-kilobase (kb) c-Ha-*ras* 1 mRNA species from A431 cells hybridizes with the exon 1–containing probe (Xma III–Bgl I), but not with the upstream-specific Bgl I–Bam HI probe. These data are consistent with our identification of this coding region as the exon nearest the 5′ end.

We have presented evidence that authentic c-Ha-*ras* 1 mRNA is initiated at four major sites between positions 537 and 561. If the region upstream from these start sites constitutes the c-Ha-*ras* 1 promoter, it should be capable of acting as a promoter in vivo. One way to demonstrate promoter activity is to insert the putative promoter sequences upstream from the bacterial chloramphenicol acetyl transferase (CAT) gene (*22*). Therefore, the 551-bp Nae I–Nae I DNA fragment (nucleotides 116 to 666), which contains about 420 bp of DNA upstream from the putative transcriptional initiation sites as well as 40 bp of the first exon and 90 bp of the first intron, was placed on the 5′ side of the CAT gene derived from the plasmic pSVOCAT (Fig. 4) (*22*). The ability of this construct (*pras*-CAT1) to express CAT activity was tested by DNA transfection. Calcium phosphate–precipitated DNA was introduced into African Green monkey CV1 cells (*23*). After 48 hours, cells transfected with *pras*-CAT 1 expressed CAT activity almost equivalent to that obtained when

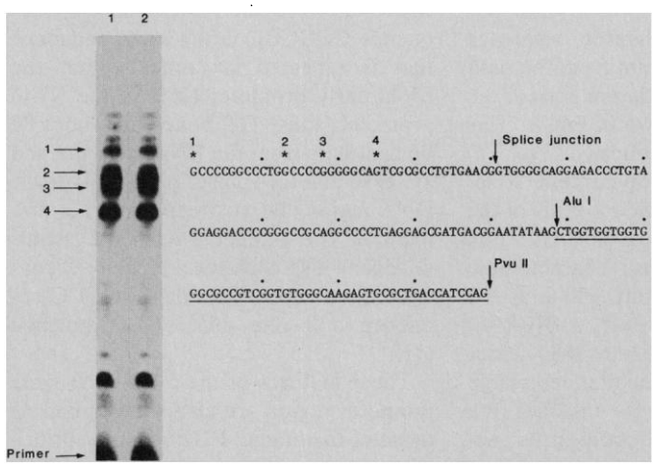
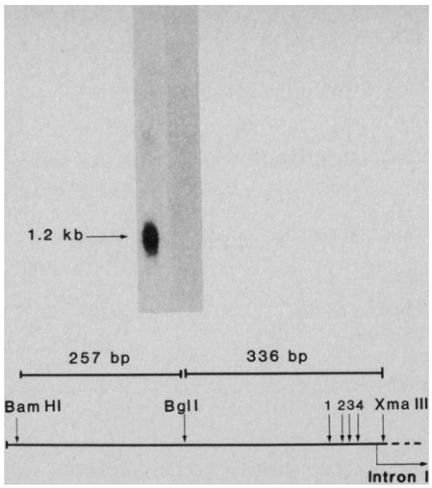


Fig. 2. Primer extension of human c-Ha-*ras* 1 mRNA and sequence of the resulting extended primers. Primer extension was performed as described (*36, 37*). The 5′ end-labeled Pvu II–Alu I DNA fragment (nucleotides 1735 to 1684) was hybridized to 5 μg of poly(A)$^+$ RNA from either A431 cells (lane 1) or pEJ6.6 - transformed NIH 3T3 cells (lane 2). Primer-extended products (arrows 1 to 4) were analyzed in a 7*M* urea–8 percent polyacrylamide gel. No products were observed in the absence of poly(A)$^+$ RNA. To determine the sequence of the extended primers, we repeated the experiment using 50 μg of A431 poly(A)$^+$ RNA. The DNA sequence of the extended primers is shown at the bottom. The numbered asterisks represent the 5′ end of each extended primer. The sequence of the 51-bp primer DNA is underlined.



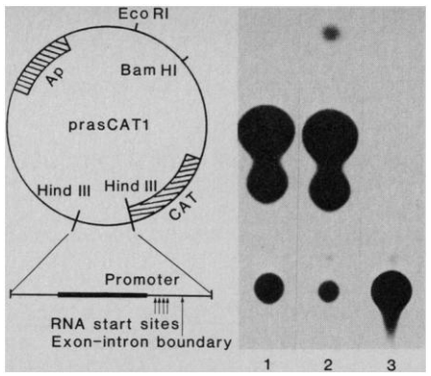


Fig. 3 (left). Northern blot analysis of c-Ha-*ras* 1 mRNA. A431 poly(A)$^+$ RNA was fractionated in 1 percent agarose containing formaldehyde and transferred to nitrocellulose (*38*). Hybridization probes were either the exon 1–containing 336-bp Xma III–Bgl I DNA fragment (nucleotides 595 to 259) (lane 1) or the upstream 257-bp Bgl I–Bam HI DNA fragment (nucleotides 259 to 2) (lane 2). Hybridization and washing were performed as described previously (*20*). Number at left indicates the size (in kilobases) of the c-Ha-*ras* 1 gene mRNA species. The partial restriction enzyme map of the 5′ untranslated region is shown at the bottom. The transcriptional initiation sites (1 to 4) are indicated. The two hybridization probes are depicted above the genomic map. Fig. 4 (right). Transient expression of the CAT gene directed by the c-Ha-*ras* 1 promoter. The structure of the p*ras*-CAT1 construct is shown at left. The diagram depicts the region of the c-Ha-*ras* 1 putative promoter region placed 5′ relative to the CAT gene (Ap, ampicillin resistance gene). The transcriptional initiation sites and the exon 1–intron 1 boundary as determined by primer extension are indicated. CV1 cells (4×10^5) were seeded on 100-mm dishes 24 hours before addition of the $CaPO_4$-DNA precipitate. One milliliter of the precipitate containing 10 μg of p*ras*-CAT1 (lane 1), pRSVCAT (lane 2), or no DNA (lane 3) was added to the medium, and the cells were incubated for 4 hours at 37°C and then exposed to glycerol for 3 minutes (*31*). Cells were harvested 48 hours after transfection, and cellular extracts were prepared by freeze-thawing three times and by sonication. After a brief centrifugation to remove cell debris, extracts were incubated with [^{14}C]chloramphenicol and acetyl coenzyme A for 30 minutes, and CAT activity was measured by thin-layer chromatography as described (*22, 23*).

the long terminal repeat (LTR) promoter of the Rous sarcoma virus (pRSVCAT) (*23*) was used (Fig. 4). Because the RSV LTR promoter was found to be very active in CV1 cells (*23*), the c-Ha-*ras* 1 promoter must be fairly strong in these cells. From these results we conclude that the 420 bp of DNA located 5′ to the observed transcriptional initiation sites of the c-Ha-*ras* 1 gene contains promoter activity.

The published nucleotide sequence (*14, 15*) of DNA upstream from the multiple transcriptional initiation sites of the c-Ha-*ras* 1 gene is shown in Fig. 1. This promoter region (nucleotides 117 to 536) has a G+C content of 80 percent. Relative to the newly identified 5′ ends of the *ras* RNA, no TATA box or CAAT box could be found in their characteristic upstream positions (about −30 and −80 bp, respectively). However, a DNA sequence at positions 362 to 368—about 170 bp upstream from the first detectable RNA start site (Fig. 1)—matches the canonical CAAT box consensus sequence (*24, 25*). We have been unable to detect an RNA start site in the region 80 bp downstream from this CAAT box–like sequence by S1 nuclease mapping. However, we cannot rule out the possibility that a minor RNA transcription start site exists here.

Further analysis of the 5′ flanking region of the c-Ha-*ras* 1 gene shows that a 6-bp sequence, the GC box CCGCCC, is repeated seven times in the promoter region at positions 157, 292, 420, 435, 512, 518, and 523. The inverted complement of this sequence, the GC box GGGCGG, occurs three times at positions 46, 341, and 401. The 6-bp sequence CCGCCC is the same sequence that is repeated six times within the SV40 early promoter (*26*). In the SV40 promoter, these GC boxes are about 80 bp upstream from the RNA start site and are essential for transcription (*27–30*). A DNA region just upstream from the first exon of the human c-Ki-*ras* 2 proto-oncogene also contains a high G+C content, three repeats of the CCGCCC sequence and one GGGCGG sequence (*31*).

These features of the c-Ha-*ras* 1 gene promoter region are also very similar to those of the human EGF receptor proto-oncogene promoter, which has a high G+C content (88 percent), contains five GC boxes, and lacks both a TATA box and a CAAT box (*16*). Other gene promoters have been described that are rich in G+C, contain multiple GC boxes, and lack the characteristic TATA and CAAT boxes, including those genes encoding hypoxanthine phosphoribosyltransferase (HPRT) (*32*), 3-hydroxy-3-methyl glutaryl coenzyme A (HMGCoA) reductase (*33*), and adenosine deaminase (*34*). Although the exact relation between the EGF receptor, c-Ha-*ras* 1, HPRT, HMGCoA reductase, and adenosine deaminase is not obvious, all are in some way involved in cellular growth control. One hypothesis is that the promoters of many of the genes involved in growth control have features similar to those described here and that regulation of their expression might involve binding of related factors. The GC boxes are present in all of these genes and may serve as attachment sites for DNA binding proteins. Gidoni *et al.* (*35*) demonstrated that the positive transcription factor purified from HeLa cells, Sp1, binds to the GC boxes of the SV40 promoter. Therefore, Sp1 or related factors may bind to these promoters and play a role in their control. Finally, the similarity between the EGF receptor gene promoter and the c-Ha-*ras* 1 gene promoter may provide a clue as to how the expression of some proto-oncogenes is regulated.

References and Notes

1. R. Ellis *et al.*, *Nature (London)* **292**, 506 (1981).
2. E. Taparowsky, K. Shimizu, M. Goldfarb, M. Wigler, *Cell* **34**, 581 (1983).
3. P. Madaule and R. Axel, *ibid.* **41**, 31 (1985).
4. C. Tabin *et al.*, *Nature (London)* **300**, 143 (1982).
5. E. Reddy, R. Reynolds, E. Santos, M. Barbacid, *ibid.*, p. 149.
6. E. Taparowsky, Y. Suard, O. Fasano, K. Shimizu, M. Goldfarb, M. Wigler, *ibid.*, p. 762.
7. K. Shimizu *et al.*, *ibid.* **304**, 497 (1983).
8. Y. Yuasa, S. Srivastava, C. Dunn, J. Rhim, E. Reddy, S. Aaronson, *ibid.* **303**, 775 (1983).
9. D. Capon *et al.*, *ibid.* **304**, 507 (1983).
10. M. C. Willingham, I. Pastan, T. Y. Shih, E. M. Scolnick, *Cell* **19**, 1005 (1980).
11. B. M. Sefton, I. S. Trowbridge, J. A. Cooper, E. M. Scolnick, *ibid.* **31**, 465 (1982).
12. E. M. Scolnick, A. G. Papageorge, T. Y. Shih, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5355 (1979).
13. R. W. Sweet, S. Yokoyama, T. Kamata, J. R. Feramisco, M. Rosenberg, M. Gross, *Nature (London)* **311**, 273 (1984).
14. D. J. Capon, E. Y. Chen, A. D. Levinson, P. H. Seeburg, D. V. Goeddel, *ibid.* **302**, 33 (1983).
15. E. P. Reddy, *Science* **220**, 1061 (1983).
16. S. Ishii, Y.-h. Xu, R. H. Stratton, B. A. Roe, G. T. Merlino, I. Pastan, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4920 (1985).
17. J. Downward, Y. Yarden, E. Mayes, G. Scrace, N. Totty, P. Stockwell, A. Ullrich, J. Schlesinger, M. D. Waterfield, *Nature (London)* **307**, 521 (1984).
18. A. Ullrich *et al.*, *ibid.* **309**, 418 (1984).
19. C. R. Lin *et al.*, *Science* **224**, 843 (1984).
20. Y.-h. Xu *et al.*, *Nature (London)* **309**, 806 (1984).
21. L. F. Parada, C. J. Tabin, C. Shih, R. A. Weinberg, *ibid.* **297**, 474 (1982).
22. C. M. Gorman, L. F. Moffat, B. H. Howard, *Mol. Cell. Biol.* **2**, 1044 (1982).
23. C. M. Gorman, G. T. Merlino, M. C. Willingham, I. Pastan, B. H. Howard, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6777 (1982).
24. C. Benoist and P. Chambon, *Nucleic Acids Res.* **8**, 127 (1980).
25. A. Efstratiadis *et al.*, *Cell* **21**, 653 (1980).
26. M. Fromm and P. Berg, *J. Mol. Appl. Genet.* **1**, 457 (1982).
27. C. Benoist and P. Chambon, *Nature (London)* **290**, 304 (1981).

28. U. Hansen and P. A. Sharp, *EMBO J.* **2**, 2293 (1983).
29. B. Wasylyk and P. Chambon, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 921 (1983).
30. R. Tjian, *Cell* **26**, 1 (1981).
31. J. P. McGrath *et al.*, *Nature (London)* **304**, 501 (1983).
32. D. W. Melton, D. S. Konecki, J. Brennand, C. T. Casky, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2147 (1984).
33. G. A. Reynolds *et al.*, *Cell* **38**, 275 (1984).
34. D. Valerio *et al.*, *EMBO J.* **4**, 437 (1985).
35. D. Gidoni, W. S. Dynan, R. Tjian, *Nature (London)* **312**, 409 (1984).
36. G. T. Merlino, J. S. Tyagi, B. deCrombrugghe, I. Pastan, *J. Biol. Chem.* **257**, 7254 (1982).
37. G. Vogeli, H. Ohkubo, M. Sobel, Y. Yamada, I. Pastan, B. deCrombrugghe, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5334 (1981).
38. P. S. Thomas, *ibid.* **77**, 5201 (1980).
39. We thank D. Lowy for providing NIH 3T3 cells transformed by pEJ6.6, J. Evans for manuscript preparation, and R. Steinberg for photography. We are grateful to R. Pozzatti and A. Johnson for reading this manuscript and for offering criticisms.

5 August 1985; accepted 30 October 1985

Sequence of the Immunodominant Epitope for the Surface Protein on Sporozoites of *Plasmodium vivax*

Abstract. Plasmodium vivax *is one of the four malaria parasites that cause disease in humans. The structure of the immunodominant repeating peptide of the circumsporozoite (CS) protein of* P. vivax *was determined. A fragment of* P. vivax *DNA that encodes this tandemly repeating epitope was isolated by use of an oligonucleotide probe whose sequence is thought to be conserved in CS protein genes. DNA sequence analysis of the* P. vivax *clone indicates that the CS repeat is nine amino acids in length (Gly-Asp-Arg-Ala-Asp-Gly-Gln-Pro-Ala). The structure of the repeating region was confirmed with synthetic peptides and monoclonal antibodies directed against* P. vivax *sporozoites. This information should allow synthesis of a vaccine for* P. vivax *that is similar to the one being tested for* P. falciparum.

THOMAS F. MCCUTCHAN
ALTAF A. LAL
VIDAL F. DE LA CRUZ
LOUIS H. MILLER
W. LEE MALOY
Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205
YUPIN CHAROENVIT
RICHARD L. BEAUDOIN
PATRICIA GUERRY
RICHARD WISTAR, JR.
Infectious Diseases Program Center, Naval Medical Research Institute, Bethesda, Maryland 20814-5055
STEPHEN L. HOFFMAN
WAYNE T. HOCKMEYER
Department of Immunology, Walter Reed Army Institute of Research, Washington, D.C. 20307
WILLIAM E. COLLINS
Malaria Branch, Division of Parasitic Diseases, Center for Infectious Diseases, Centers for Disease Control, Atlanta, Georgia 30333
DYANN WIRTH
Department of Tropical Public Health, Harvard School of Public Health, Boston, Massachusetts 02114

Of the four human malarias, those due to the parasites *Plasmodium vivax* and *Plasmodium falciparum* are the most common and are a major cause of disease in the tropics. Controlling these two infections would improve health in this region. One research approach has been the development of vaccines against the parasite at different stages in its life cycle (*1*). Sporozoites, the form inoculated by mosquitoes to initiate the human infection, are covered with a protein known as the circumsporozoite (CS) protein (*2*). Antibodies to the CS protein block infection in vivo (*3*). The gene for the CS protein of *P. falciparum* has been cloned by using an antibody to the CS protein to screen a genomic library constructed with mung bean nuclease–digested fragments (*4*) or a complementary DNA (cDNA) library from sporozoite messenger RNA (*5*). Recombinant (*6*) and synthetic peptide antigens (*7*) of the repeat region in the middle of the CS protein of *P. falciparum* induce antibodies that block sporozoite invasion of liver cells in vitro. These antigens will be used for vaccine testing in man. We now report the sequence of the immunodominant repeat region of the CS protein of *P. vivax*.

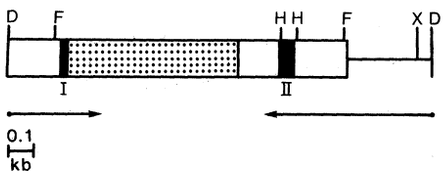


Fig. 1. General structure and partial restriction map of the CS protein gene of *P. vivax*. DNA sequence analysis (arrows) established the orientation of the gene (box) (the amino-terminus is to the left, the carboxyl-terminus to the right) and yielded the nucleotide sequence of the repeating epitope as shown in Fig. 2. Regions I and II are indicated by black boxes. The repeat region is indicated by stippling. Restriction enzyme sites indicated: D, Dra I; F, Fok I; H, Hinf I; X, Xba I.

Our strategy for cloning the gene for the CS protein of *P. vivax* with an oligonucleotide probe from a region of nucleotide homology between *P. falciparum* and *P. knowlesi* (region II) (*4*) was based on the observation (*8*) that *P. vivax* is more closely related to *P. knowlesi* than to *P. falciparum*. Therefore, if a sequence is conserved in *P. knowlesi* and *P. falciparum*, it should also be conserved in *P. vivax*. A 27-base oligonucleotide probe to region II (*9*) hybridized to a single fragment 1.3 kilobases in length of Dra I–digested genomic DNA (*10*) of an El Salvador strain (Sal-I) of *P. vivax* (*11*). The same probe was then used to screen a Dra I genomic library cloned in the bacterial plasmid pUC9.

The general structure of the CS protein gene from *P. vivax* as deduced from the pPv1 insert is shown in Fig. 1. The nucleotide sequence of a segment of the gene containing the repeats of the CS protein of *P. vivax* is shown in Fig. 2. It contains tandem repeats of nine amino acids (Gly-Asp-Arg-Ala-Asp-Gly-Gln-Pro-Ala). The peptide on the amino-terminal side of the repeat region has an overall structure similar to those of the CS proteins of *P. knowlesi* (*12*) and *P. falciparum* (*4*), including a charged region and the homologous region I (Fig. 3). Because of the high degree of homology between *P. vivax* and *P. knowlesi* (see below), we predicted that the 5′ coding region of the gene missing from our Dra I clone would encode the start codon for translation, a signal sequence, and a few amino acids between the signal sequence and the Dra I recognition sequence.

Monoclonal antibodies against sporozoites of *P. knowlesi*, *P. falciparum*, and *P. cynomolgi* reacted with tandemly repeated peptides in the CS proteins (*4, 5, 12, 13*). Therefore, to prove that the sequence was indeed the CS protein of *P. vivax*, we investigated how monoclonal antibodies to *P. vivax* sporozoites would react to a synthetic peptide of two tandem repeats (*14*) in the deduced *P. vivax* CS peptide. The peptide (at 100 μg/ml) blocked the binding of two different horseradish peroxidase–labeled monoclonal antibodies (219c and 427) to *P. vivax* sporozoites in an enzyme-linked immunosorbent assay. As a control for specificity, we determined the effect of the repeat peptide (four repeats) of the CS protein of *P. falciparum* on the binding of these two antibodies to *P. vivax* sporozoites. The *P. falciparum* peptide did not block the binding of antibody to *P. vivax* sporozoites. Furthermore, antibody 219c reacted with the *P. vivax* peptide in a two-sided immunoassay