Volume 302, number 3, 269–273 FEBS 11043 May 1992

Structural analysis of the putative regulatory region of the rat gene encoding poly(ADP-ribose) polymerase

Frédéric Potvin*, Jacques Thibodeau*, James B. Kirkland, Brigitte Dandenault, Caroline Duchaine and Guy G. Poirier

Laboratoire du métabolisme du poly(ADP-ribose), Endocrinologie moléculaire, Centre Hospitalier de l'Université Laval, 2705 Boul. Laurier, Ste-Foy, Québec G1V 4G2, Canada

Received 19 February 1992

A λ EMBL3 clone containing the first three exons along with part of the 4th exon of the rat poly(ADP-ribose) polymerase gene was isolated from a genomic DNA library. This clone also contains 6.6 kbp of upstream sequences. Nucleotide sequence analysis of the proximal 5′ 670 nucleotides flanking the major RNA start site of the rat gene does not reveal significant global homology with the same region of the human gene, but a series of short sequences are identical. Among these sequences are found two putative Sp1 binding sites along with a decanucleotide sequence responsible for the attachment of the transcription factor AP-2.

Poly(ADP-ribose) polymerase; Genomic DNA cloning; Nucleotide sequence; Regulatory region; Sp1

1. INTRODUCTION

Poly(ADP-ribose) polymerase (PARP) is a 113 kDa DNA-binding enzyme found in the nucleus of almost all eukaryotic cells [1]. Its activation is dependent on the presence of DNA strand breaks [2], and results in the synthesis of a homopolymer of ADP-ribose units on a variety of nuclear proteins, including histones, DNA polymerases and PARP itself [3,4]. Synthesis of poly(ADP-ribose) on histones causes relaxation of chromatin structure [1], and the enzyme is thought to play a role in DNA repair, replication recombination and expression.

It has been suggested that phorbol esters can stimulate de novo PARP mRNA and protein synthesis in fibroblasts [5], implying that this product might have an inducing effect on the PARP gene promoter.

The human 43 kbp gene has been isolated and consists of 23 exons [6] located in the q42 region of chromosome 1 [7,8]. The promoter of the human gene shows structural similarities to that of the DNA polymerase β gene [9]. Both of these genes are involved in DNA repair, but the synthesis of DNA polymerase β is induced by DNA damage [10] while that of the PARP gene is not [11]. The putative functional features of the human PARP promoter consist of 2 Sp1 binding sites separated by a 26 nucleotide palindromic sequence [9] and 3 AP-2 recognition sequences [12]. As in many housekeeping genes [13], both the DNA polymerase β gene [14] and the PARP gene lack typical TATA and CCAAT sequence elements [9,12]. Yokoyama et al. demonstrated that the human PARP gene promoter is inducible by cAMP and phorbol esters [12]. Although the coding sequence of the PARP mRNAs are well-conserved between human [8,15,16], mouse [17], rat [18], bovine [19], and chicken [20], comparative studies on the structure of the regulatory regions of the gene have not been reported. In order to study the regulation of the PARP gene, we have undertaken the cloning of the promoter region of the rat gene. Here we report the isolation of a λ genomic clone containing 6.6 kbp of 5′ flanking sequences and the first 4 exons of the rat PARP gene. We also report the sequence from nucleotide −670 to +143.

Abbreviations: PARP, poly(ADP-ribose) polymerase; kDa, kilo Daltons; kbp, kilo base pair(s); cAMP, cyclic AMP.

Correspondence address: G. Poirier, Centre Hospitalier de l'Université Laval, Laboratoire du métabolisme du poly(ADP-ribose), Endocrinologie moléculaire, 2705 Boul. Laurier, Ste-Foy, Que. G1V 4G2, Canada. Fax: (1) (418) 654-2761.

*Both authors contributed equally to the paper.

2. MATERIALS AND METHODS

2.1. *DNA probes and genomic library screening*

Two cDNA probes were used in this study. The first consists of the 5′ 700 bp coding sequence of the human PARP cDNA and was described previously [21]. This insert was excised from vector sequences using *Pst*I and *Hin*dIII and used to locate the exons of the gene on the different restriction fragments. The second probe is the 125 bp 5′ portion of the same insert and was excised using *Pst*I and *Kpn*I. This smaller fragment extends from the first ATG codon into the coding region of the PARP cDNA and was used for the screening of the library. Both fragments were purified by low-melting point agarose gel electrophoresis [22] (BRL Ultra-pure Reagents) and labeled with [^{32}P]

to specific activities of about 10^9 cpm/μg using Amersham's Random Priming Kit.

The Sprague–Dawley rat genomic library, constructed in phage λ EMBL3 was described previously [23]. The probe used for screening was the 125 bp *Pst*I–*Kpn*I fragment. A total of 2×10^6 phage plaques were screened in duplicate according to the procedure of Davis et al. [24]. Washes were in 2× SSC at 60°C, and the autoradiogram was exposed 2 h at −80°C with intensifying screens.

2.2. *Restriction mapping, Southern blots and DNA sequencing*

Lambda phage DNA was prepared as described [24] and the cutting sites of the chosen restriction enzymes (Pharmacia) were determined by 1% agarose gel field inversion electrophoresis [25,26]. For bi-directional Southern blotting experiments [27], λ phage DNA was cut with restriction enzymes and separated on 1% agarose gels. DNA was transferred onto nylon membranes (Amersham), hybridized [28] and washed as described above. The genomic Southern blot (Zoo blot) was purchased from Clonetech, hybridized and washed as above. The rat DNA fragments of interest were subcloned from phage λ into Bluescript vectors (Stratagene). Deletions using exonuclease III (Pharmacia) and dideoxy-sequencing (USB) of both strands were performed as described [18].

2.3. *Transcription start site mapping*

Transcription start sites were determined by the Nuclease S1 mapping technique [22]. The probe used consisted of the *Xba*I–*Eco*RI restriction fragment of the 800 bp subclone (construction, Fig. 1D; sequence, Fig. 2) which was end-labeled with T4 polynucleotide kinase (Pharmacia) to a specific activity of 2×10^6 cpm/μg. This probe was denatured and hybridized with rat thymus mRNA. The DNA–RNA hybrids were then digested with S1 nuclease (Pharmacia) and sized by polyacrylamide-urea gel electrophoresis (BRL Ultra-pure Reagents) along with the sequence reaction of the original clone (Fig. 3).

3. RESULTS AND DISCUSSION

In order to isolate the regulatory region of the PARP gene, we screened a rat genomic DNA library with the first 125 coding bp of the human cDNA. The specificity of this probe was assessed by hybridizing it with genomic DNA from different organisms digested with *Eco*RI (Fig. 1A). Under the conditions used for hybridization the probe is very specific, as only one band of about 15 kbp hybridizes in the rat DNA (Fig. 1A, lane 1). The present of only one band contrasts with the results obtained on human DNA using a similar probe [8] and suggests the presence of only one PARP gene in rat. This one-band hybridization pattern is also found

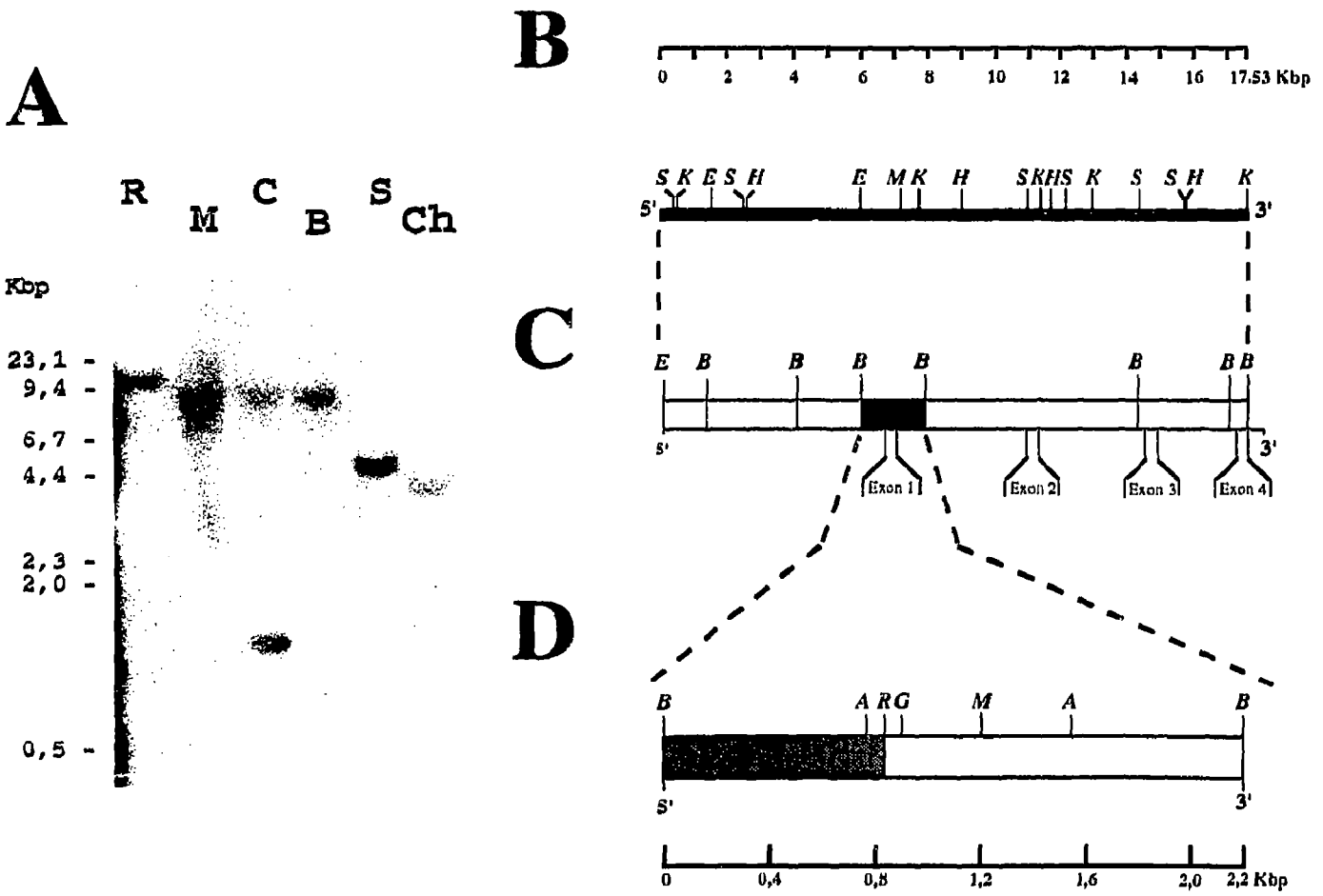

Fig. 1. Organization of pGRat5'. (A) Southern blot analysis of the structure of the 5' portion of the PARP gene in different organisms using the 125 bp 5' human cDNA probe. Molecular weight markers are in kbp. R, rat; M, mouse; C, canine; B, bovine; S, sheep; Ch, chicken. (B) Restriction map of pGRat5'. (C) Proposed organization of the PARP sequences within pGRat'. (D) Fine restriction mapping of the upstream *Bam*HI-hybridizing fragment. The shadowed *Rsa*I–*Bam*HI fragment has been sequenced. S, *Sac*I; K, *Kpn*I; E, *Eco*RI; H, *Hin*dIII; M, *Sma*I; B, *Bam*HI; A, *Ava*II; R, *Rsa*I; G, *Bgl*II.

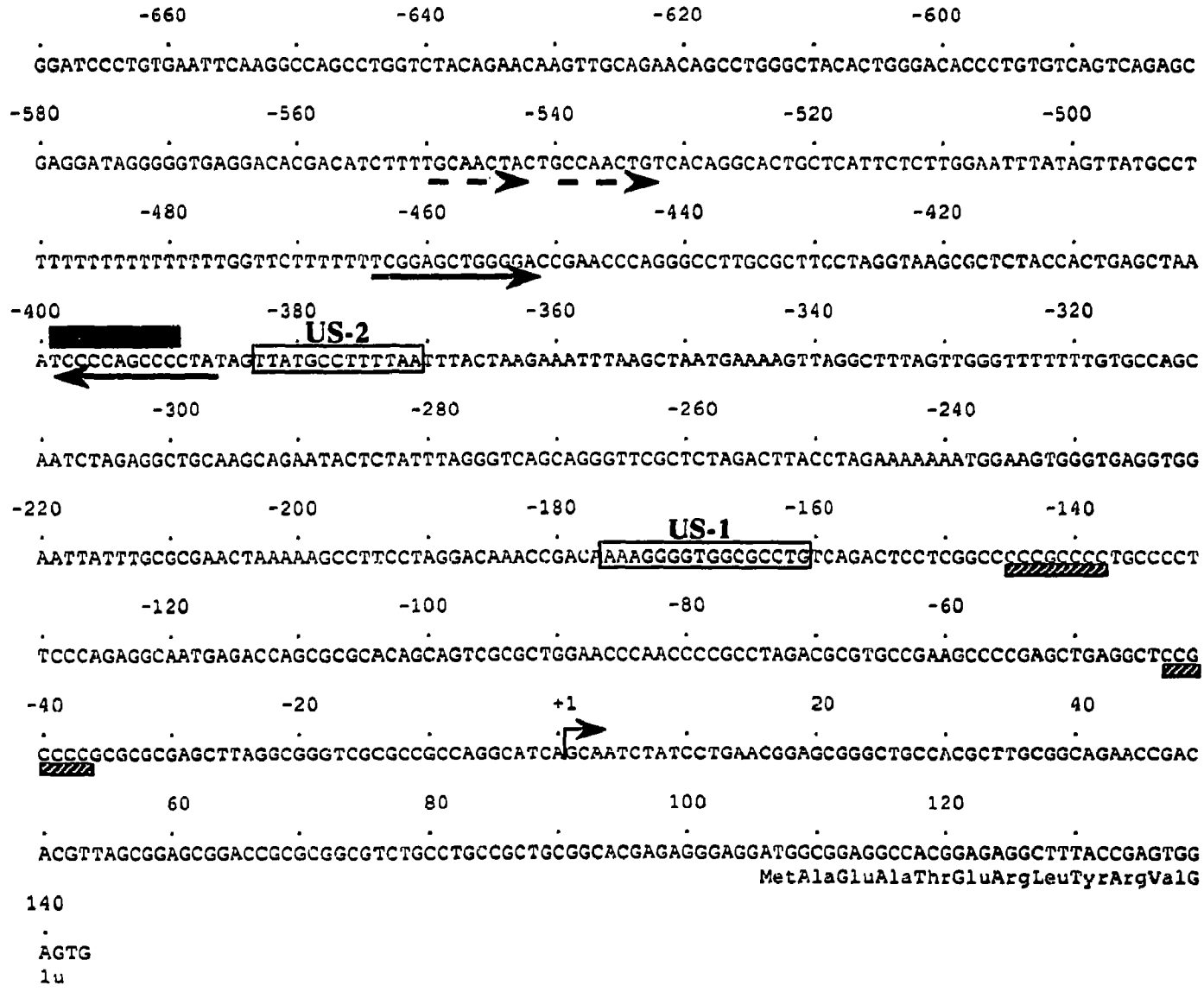

Fig. 2. Sequence of the upstream region of the rat PARP gene described from −670 to +143. Boxed regions indicate US-1 and US-2 elements. The broken arrow represents the major transcription initiation site. Thick dashed arrows show direct repeats, and thick plain arrows represent inverted repeats. The solid black box indicates the location of the AP-2 recognition site and striped boxes show the location of GC boxes.

with mouse, bovine, sheep and chicken DNA (Fig. 1A, lanes 2,4,5 and 6, respectively). About 2×10^6 recombinant plaques were screened in the same conditions, and only one hybridization-positive clone was isolated. This clone was called pGRat5′ and contained 17.5 kbp of rat sequences.

The restriction map of this insert was determined (Fig. 1B) and the results are consistent with the Southern blotting experiments on the genomic rat DNA as the insert contains an *Eco*RI fragment of at least 10 kbp which hybridized with cDNA probes (see above). To determine which other fragments of pGRat5′ were homologous with PARP cDNA we hybridized *Bam*HI digests of pGRat5′ DNA with the 125 bp and 700 bp probes (not shown). Four *Bam*HI fragments (Fig. 1C) hybridized with the 700 bp probe. Two of these fragments (2.2 and 6.4 kbp) are contiguous and hybridized only with the shorter 5′ 125 bp probe. This result allowed us to determine the 5′ and 3′ ends of the insert. The sequence of the last 38 nucleotides of the most 3′ *Bam*HI fragment showed 95% homology with the fourth exon of the human PARP gene. The 5′ splice site of this exon would also be conserved between the two genes (not shown). The approximate position of exons 2 and 3 was determined by hybridizing restriction digests of the *Bam*HI subclones bearing these exons with the 700 bp probe. By using several different restriction enzymes we have been able to ascertain their position (Fig. 1C). We then sequenced a *Bam*HI–*Rsa*I fragment from the extreme 5′ hybridizing *Bam*HI fragment, and

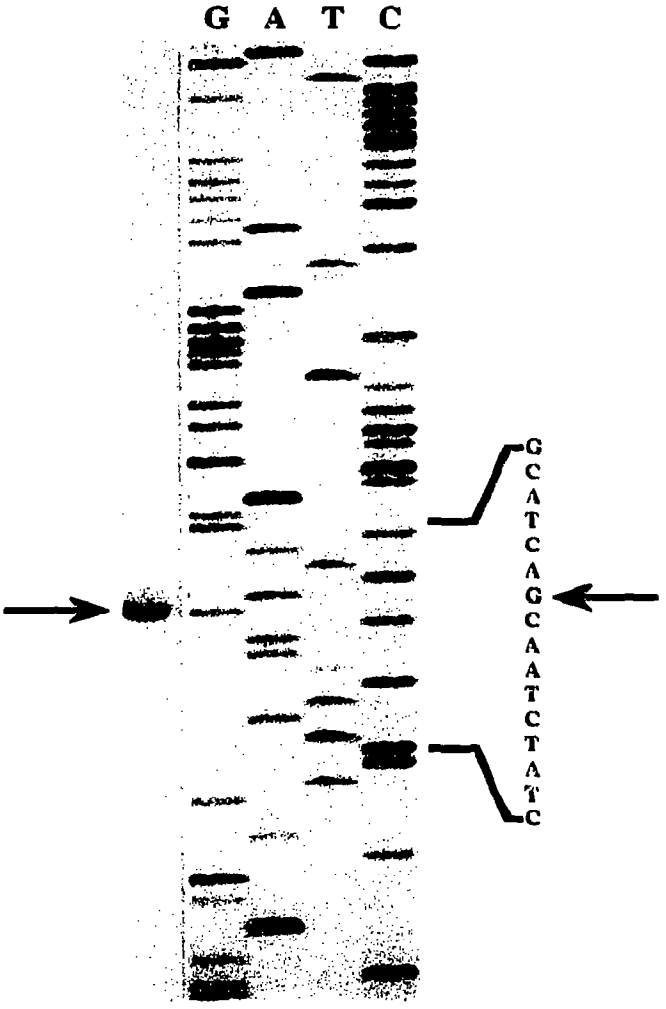

Fig. 3. Localization of the transcription start sites. The results of the S1 mapping experiment appear at the left (arrow) along with the corresponding sequence at the right. The probe used was an *Hin*fI–*Eco*RI digest of the clone under study (Fig. 2) and corresponded to nucleotide −157 to +413.

the nucleotide sequence is shown in Fig. 2. By homology with the sequence of the human cDNA we have identified the first ATG codon. The approximate position of exons 2 and 3 on the internal *Bam*HI hybridizing fragments is shown in Fig. 2. The deduced amino acid sequence encoded by part of the first exon (Fig. 2) is 98% homologous to the same region of the human protein sequence [16]. The sequences upstream of the first ATG in the rat and the human PARP gene [9] were compared but no overall homology was found. However, both regions contain GC-rich boxes (CCGCCCC) which are putative Sp1 binding sites, known to be essential for the activity of many other promoters [29,30] (Fig. 2). The sequence between the first translated ATG and the most upstream GC boxes (Fig. 2) is GC rich (69%). It also contains many CpG dinucleotides, which are putative methylation sites and are often involved in the tissue-specific regulation of gene expression [31,32].

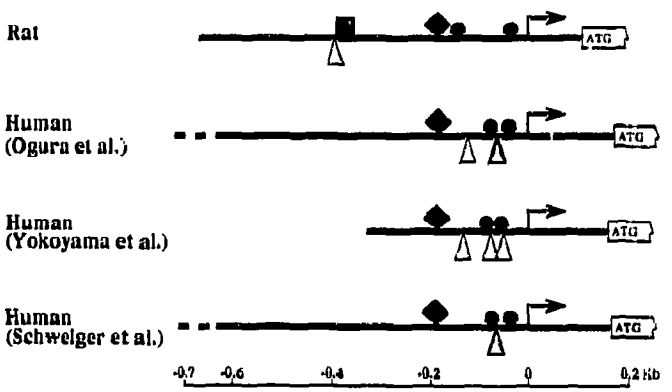

Fig. 4. Comparison of the organization of the promoters of rat and human PARP genes. The arrows indicate major initiation sites, the filled circles represent the GC boxes and the open triangles are for AP-2 recognition sites. The US-1 and US-2 elements are represented, respectively, by shaded diamonds and filled squares. The US-2 element is located far upstream on human clones (not shown).

In the human gene, Ogura et al. [9] identified a 14 nucleotide region located 171 bp upstream of the first ATG which contains 10 transcription starts. This 14 nucleotide region, with only one substitution, is also found in the rat DNA downstream of the GC boxes (Fig. 2). The major transcription initiation site in the rat is also located in this region, as determined by S1 mapping experiments (Fig. 3). Overexposure of the autoradiogram shows the appearance of a second initiation size aligning on the next nucleotide on the sequence (data not shown). The same pattern followed by a few very weak bands have been revealed by primer extension (not shown). This creates a mRNA 5′ leader sequence about 60 bp shorter than the average human leader (Fig. 4) and would be consistent with the difference in size observed between the rat and human mRNAs [18].

We have also found in pGRat5′ a 10 bp sequence, at position −390, which corresponds to the recognition site of the transcription factor AP-2. It has been shown that this element confers the inducibility of a gene by cAMP and phorbol esters [12]. It is likely to be active in the environment of the PARP promoter. Indeed, it has been demonstrated by Yokoyama et al. that the expression of the human PARP gene is stimulated by these two compounds [12], and suggested by Singh et al. that phorbol esters promote de novo PARP mRNA synthesis [5]. All of the *cis*-acting regulatory sequences described by Jones et al. [33] were compared to the sequence pGRat5′, but the only homologous sequences found were the Sp1 sites and the AP-2 consensus sequence [34].

Moreover, we did find two short sequences (Upstream Sequences 1 and 2, US-1 and US-2 in Fig. 2) which are present with high homology in the human PARP gene [9,12,38], and with some mismatches in DNA polymerase β genes from the rat [35], mouse [35],

and human [14]. These sequences (US-1 and US-2) show no homology with any consensus sequence recognized by a known DNA-binding protein or transcription factor [33,37]. Furthermore, they have been found in the sequences of the many housekeeping genes we have investigated, from polymerase β to histones.

The general organization of the rat PARP gene promoter reported here is quite similar to the three human sequences published so far [9,12,38] (Fig. 4). Indeed, all these sequences bear GC boxes, putative AP-2 recognition sequences, and the two 'upstream sequences'. The two GC boxes are located between −37 and −145 in all the promoters. All the promoters also bear putative AP-2 recognition sites, although they vary in number and position. Interestingly, the US-1 appears to be highly conserved and constant in its position. This element could be essential for the basal expression of the PARP gene. The US-2 element is less conserved and its position differs between rat and human sequences [9,38]. This element might act on the PARP gene as an enhancer, although the effects of both 'upstream sequences' on the expression level of the PARP gene are yet to be ascertained.

Acknowledgements: We thank Dr. M. Guertin for the gift of the λ EMBL3 library, Dr. G. Gradwhol for the gift of the pGG700 probe and M.C. Trudel for providing the rat RNA. We also want to thank J. Poulin, A. Douville and L. Doyon for excellent secretarial assistance. J.T. is supported by La Société de la Recherche sur le Cancer Inc. We also thank Dr. S. Guérin for helpful discussions and Dr. Girish Shah for reviewing this paper. This work was supported by the MRC.

REFERENCES

[1] de Murcia, G., Huletsky, A. and Poirier, G.G. (1988) Biochem. Cell Biol. 66, 626–635.
[2] Benjamin, R.C. and Gill, D.M. (1980) J. Biol. Chem. 255, 10502–10508.
[3] Althaus, F.R. and Ritcher, C.R. (1987) Mol. Biol. Biochem. Biophys. 37, 1–125.
[4] Ueda, K. and Hayaishi, O. (1985) Annu. Rev. Biochem. 54, 73–100.
[5] Singh, N., Poirier, G.G. and Cerutti, P. (1985) EMBO J. 4, 1491–1494.
[6] Auer, B., Nagl, U., Herzog, H., Schneider, R. and Schweiger, M. (1989) DNA 8, 575–580.
[7] Herzog, H., Zadel, B.U., Schneider, R., Auer, B., Hirsch-Kauffman, M. and Schweiger, M. (1989) Proc. Natl. Acad. Sci. USA 86, 3514–3518.
[8] Cherney, B.W., McBride, W.W., Chen, O., Alkhatib, H., Bathia, K., Hensely, P. and Smulson, M.E. (1987) Proc. Natl. Acad. Sci. USA 84, 8370–8374.
[9] Ogura, T., Nyunoya, H., Takahashi-Masutani, M., Miwa, M., Sugimura, T. and Esumi, H. (1990) Biochem. Biophys. Res. Commun. 167, 701–710.
[10] Fornace Jr., A.J., Zmudzka, B., Hollander, M.C. and Wilson, S.H. (1989) Mol. Cell Biol. 9, 851–853.
[11] Bhatia, K., Pommier, Y., Giri, C., Fornace, A.J., Imaizumi, M., Breitman, T.R., Cherney, B.W. and Smulson, M.E. (1990) Carcinogenesis 11, 123–128.
[12] Yokoyama, Y., Kawamoto, T., Mitsuuchi, Y., Kurosaki, T., Toda, K., Ushiro, H., Terashima, M., Sumimoto, H., Kuribayashi, I., Yamamoto, Y., Maeda, T., Ikeda, H., Sagara, Y. and Shizuta, Y. (1990) Eur. J. Biochem. 194, 521–526.
[13] Dynan, W.S. (1986) Trends Genet. 2, 196–197.
[14] Widen, S.G., Kedar, P. and Wilson, S.H. (1988) J. Biol. Chem. 263, 16992–16998.
[15] Kurosaki, T., Ushiro, H., Mitsuruchi, Y., Suzuki, S., Matsuda, N., Matsuda, Y., Katunuma, N., Kangawa, K., Matsuo, H., Hirose, T., Inayama, S. and Shizuta, Y. (1987) J. Biol. Chem. 262, 15990–15997.
[16] Uchida, K., Morita, T., Sato, T., Ogura, T., Yamashita, R., Noguchi, S., Suzuki, H., Yunoya, H., Miwa, M. and Sugimura, T. (1987) Biochem. Biophys. Res. Commun. 148, 617–622.
[17] Huppi, K., Bhatia, K., Siwarski, D., Klinman, D., Cherney, B. and Smulson, M. (1989) Nucleic Acids Res. 17, 3387–3401.
[18] Thibodeau, J., Gradwhol, G., Dumas, C., Clairoux-Moreau, S., Brunet, G., Penning, C., Poirier, G.G. and Moreau, P. (1989) Biochem. Cell Biol. 67, 653–660.
[19] Taniguchi, T., Yamauchi, K., Yamamoto, T., Yoyoshima, K., Harada, N., Tanaka, N., Takahash, S., Yamamoto, H. and Fugimoto, S. (1988) Eur. J. Biochem. 171, 571–575.
[20] Ittel, M.E., Garnier, J.M., Jeltsh, J.M. and Niedergang, C.P. (1991) Gene 102, 157–164.
[21] Gradwohl, G., Ménissier-de Murcia, J., Molinette, M., Simonin, F., Koken, F., Hoeijmakers, J.H.J. and de Murcia, G. (1990) Proc. Natl. Acad. Sci. USA 87, 2990–2994.
[22] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor.
[23] Chevrette, M., Guertin, M., Turcotte, B. and Bélanger, L. (1987) Nucleic Acids Res. 15, 1338–1339.
[24] Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) Basic Methods in Molecular Biology, Elsevier, New York.
[25] Birren, B.W., Lai, E., Hood, L. and Simon, M.I. (1989) Anal. Biochem. 177, 282–286.
[26] Rackwitz, H.R., Zehetner, G., Frischauf, A.M. and Lehrach, H. (1984) Gene 30, 195–200.
[27] Smith, G.E. and Summers, M.D. (1980) Anal. Biochem. 109, 123–129.
[28] Meinkoth, J. and Wahl, G. (1984) Anal. Biochem. 138, 267–284.
[29] Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1987) Cell 51, 1079–1090.
[30] Kadonaga, J.T., Carner, K.R., Masiarz, F.R. and Tjian, R. (1987) Cell 51, 1079–1090.
[31] Cedar, H. (1988) Cell 53, 3–4.
[32] Dynan, W.S. and Tjian, R. (1989) Trends Genet. 5, 35–36.
[33] Jones, N.C., Rigby, P.W.J. and Ziff, E.B. (1988) Gene Dev. 2, 267–281.
[34] Imagawa, M., Chiu, R. and Karin, M. (1987) Cell 51, 251–260.
[35] Yamagushi, M., Hirose, F., Hayash, Y., Nishimoto, Y. and Matsukage, A. (1987) Mol. Cell Biol. 7, 2012–2018.
[36] Reynolds, G.A., Basu, S.K., Osborne, T.F., Chin, D.J., Gil, G., Brown, M.S., Goldstein, J.L. and Luskey, L. (1984) Cell 278, 275–285.
[37] Mitchell, P.J. and Tjian, R. (1989) Science 245, 371–378.
[38] Schweiger, M., Auer, B., Herzog, H., Hirsch-Kauffmann, M., Kaiser, P., Flick, K., Nagl, U. and Schneider, R. (1992) In: ADP-Ribosylation reactions (Poirier and Moreau, eds.) Springer Verlag, New York. Genebank accession number #M60436 (in press).