Comparative Structural Analysis of the Transcriptionally Active Proopiomelanocortin Genes A and B of *Xenopus laevis*¹

Peter M. T. Deen, Marion J. G. Bussemakers, Dick Terwel, Eric W. Roubos, and Gerard J. M. Martens

Department of Animal Physiology, University of Nijmegen

In the intermediate lobe of the pituitary gland, the prohormone proopiomelanocortin (POMC) is processed to, among other peptides, melanocyte-stimulating hormone (α-MSH). In the toad *Xenopus laevis* α-MSH controls skin darkening during background adaptation, and the level of POMC gene transcription in the intermediate lobe depends on the color of the background. In the lobe, two structurally different POMC proteins are produced from two mRNAs that are transcribed to approximately the same level from two POMC genes (A and B). We previously reported the entire nucleotide sequence of *Xenopus* POMC gene B. To identify conserved—and thus potential regulatory—DNA elements in the *Xenopus* POMC gene, we here report the determination and analysis of the complete nucleotide sequence of *Xenopus* POMC gene A and its 5′- and 3′-flanking regions. Comparison of the two *Xenopus* POMC genes revealed, in addition to the exons, three highly conserved regions. First, the promoter regions are >90% identical. The second region concerns JH12 repetitive elements situated at approximately the same position in both genes. These elements are >86% identical. The third region is a 500-bp sequence just upstream of exon three (63% identity). Besides these three large regions, several small regions with significant identity were found at similar positions in the two POMC genes. The fact that, except for the JH12 element, the repetitive elements are not conserved between the two POMC genes indicates that these repeats are not functionally important.

Introduction

One of the aims of our research is to get a better understanding of the activation and inactivation of genes that are responsible for a proper functioning of neuroendocrine cells. For this purpose we are studying, as a model, the proopiomelanocortin (POMC) gene in the intermediate lobe of the pituitary of the South African clawed toad, *Xenopus laevis*. The POMC gene is predominantly expressed in the pituitary, where the precursor protein is tissue-specifically processed. In the corticotrope cells of the anterior pituitary, the end products of POMC processing are adrenocorticotropic hormone (ACTH), while in the melanotrope cells of the neurointermediate lobe ACTH is further processed to corticotropin-like intermediate-lobe peptide (CLIP) and melanocyte-stimulating hormone (α-MSH) (for review, see Smith and Funder)

¹ Key words: proopiomelanocortin gene, *Xenopus*, sequence identity, regulatory DNA elements.

² Abbreviations: POMC = proopiomelanocortin; MSH = melanocyte-stimulating hormone; ACTH = adrenocorticotropic hormone; Vi = vitellogenin.

Address for correspondence and reprints: Peter M. T. Deen, Department of Animal Physiology, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands.

Mol. Biol. Evol. 9(3):483-494, 1992. © 1992 by The University of Chicago. All rights reserved.
0737-4038/92/0903-0009$02.00
In amphibians α-MSH mediates background adaptation (for review, see Bagnara and Hadley 1973). Placing a toad on a black background results, via a neuroendocrine reflex, in α-MSH release from the melanotrope cells. α-MSH induces dispersion of pigment granules in skin melanophores, which is followed by a darkening of the toad. On a white background α-MSH release is inhibited, resulting in aggregation of the pigment granules and thus a pale toad.

An elevated release of α-MSH by melanotrope cells of black animals is accompanied by a higher rate of POMC gene transcription (Martens et al. 1987; T. A. Y. Ayoubi, personal communication) and POMC biosynthesis (Verburg-van Kemenade et al. 1984; Loh et al. 1985). In black background–adapted toads the level of POMC mRNA in the intermediate lobe is at least 15 times higher than that in white-adapted animals (Martens et al. 1987). Biosynthetic studies have revealed that two different POMC molecules are synthesized in the intermediate lobe in similar amounts (Martens et al. 1982). Therefore, in *X. laevis* α-MSH is a cleavage product of two structurally different POMC precursors. In view of the numbers of cDNA clones encoding POMC-A and -B, both proteins appear to be translated from virtually equal amounts of POMC mRNAs A and B (Martens 1986), indicating that two POMC genes are expressed to approximately the same level in the melanotrope cells of the *Xenopus* intermediate lobe. This observation, together with the amount of POMC mRNA being much higher in melanotrope cells of black-adapted toads than in those of white-adapted animals, indicates that POMC genes A and B are coordinately regulated in these cells. The phenomenon of coexpressed genes led us to investigate the complete structures of the two POMC genes in order to identify conserved—and thus potential regulatory—regions in the genes.

POMC mRNA-B of *Xenopus* is encoded by the previously reported POMC gene (Martens 1987; Deen et al. 1991). This *Xenopus* POMC gene B exhibits essentially the same structural organization as do mammalian POMC genes—namely, two introns (A and B) separating three exons, the third of which encodes all bioactive POMC-derived peptides (Nakanishi et al. 1979; Takahashi et al. 1981; Drouin et al. 1985). Here we report on the structural organization of POMC gene A as well as on a comparative analysis of *Xenopus* POMC genes A and B that reveals conserved nucleotide sequences that may be involved in the regulation of POMC gene transcription.

**Material and Methods**

**Isolation and Restriction Mapping of *Xenopus* POMC Gene A**

Clone λXPA5 was isolated from a *X. laevis* genomic library (provided by Dr. I. Dawid, National Institutes of Health, Bethesda, Md.). This clone is a λEMBL4 vector with a 13.2-kb *Xenopus* genomic DNA fragment containing POMC gene A. Restriction mapping of λXPA5 was carried out by Southern blot hybridization analysis using as probes cDNA clones pXPL, pXP20, and pXP123, together covering both part of the 5'- and 3'-untranslated mRNA regions and the complete protein-encoding mRNA region (Martens 1987). The cDNA clones were labeled with [α-32P]dATP by nick-translation (Weinstock et al. 1978) or random priming (Feinberg and Vogelstein 1983). For further restriction mapping and sequence analysis, restriction fragments of λXPA5 DNA were subcloned into PBR322 and pUC18/19.

**Identification of Repetitive Elements in *Xenopus* POMC Gene A**

To locate repetitive elements in *Xenopus* POMC gene A, we hybridized Southern blots of *Xenopus* POMC gene fragments, obtained by digestion with several endo-
nucleases, with a probe derived from *Xenopus* genomic DNA according to a method described elsewhere (Deen et al. 1991).

**DNA Sequence Analysis**

DNA sequencing was performed by the dideoxy chain-termination method (Sanger et al. 1977) using either (a) M13 mp10/11, mp18/19, pKUN19, or PUC18/19 subclones as templates or (b) M13mp10/11 deletion mutants generated after limited cleavage of the insert DNA by DNaseI in the presence of Mn2+, according to Hong's (1982) method as modified by Lin et al. (1985). About 90% of the nucleotide sequence was determined on both strands.

**Computer Sequence Analysis**

Similarity searches between nucleotide sequences were performed with the computer facility of the University of Nijmegen, which uses the algorithms of Wilbur and Lipman (1983). For identification of repetitive elements in both the *Xenopus* POMC gene and its flanking regions, the EMBL (release 24) and GenBank (release 60.0) data bases were used.

**Results and Discussion**

The Entire Nucleotide Sequence of *Xenopus* POMC Gene A

Screening of a λEMBL4 *X. laevis* genomic library with POMC cDNA clone pXP20, which contains the nearly full-length nucleotide sequence of *Xenopus* POMC gene transcript B (Martens 1986), resulted in the isolation of 37 hybridization-positive clones. Rescreening of these clones with a probe containing the first exon of *Xenopus* POMC gene B (MλXP5-248) gave eight positives. On detailed restriction mapping and hybridization of the restriction fragments with probes pXP20 and MλXP5-248, it became clear that seven of these clones contained *Xenopus* POMC gene B, while one clone (λXP5A) gave completely different restriction patterns. Both sequence analysis of λXP5A DNA fragments hybridizing with the POMC cDNA probes and comparison of these nucleotide sequences with a cDNA corresponding to POMC gene transcript A (Martens 1986) showed that the 13.2-kb insert of clone λXP5A contained the entire *Xenopus* POMC gene A. The finding that two POMC genes occur in *X. laevis* is not surprising, since the genome of this species appears to have arisen by chromosome duplication (tetraploidization) that happened some 30 Mya (Bisbee et al. 1977; Thiebaud and Fischberg 1977).

The nucleotide sequence of *Xenopus* POMC gene A and its 5'- and 3'-flanking regions is given in figure 1. From this figure we conclude that gene A has the same structural organization as does the previously described gene B (Martens 1987; Deen et al. 1991). Like their mammalian counterparts (Nakanishi et al. 1979; Takahashi et al. 1981; Drouin et al. 1985), the *Xenopus* POMC genes contain two introns (A and B) separating three exons, of which the third exon encodes all bioactive peptides. The length of *Xenopus* POMC gene A is 8,432 bp, which is significantly larger than *Xenopus* POMC gene B (6,030 bp; Deen et al. 1991). The lengths of exon 1 (48 bp), exon 2 (143 bp), and exon 3 (1,012 bp) of POMC gene A are identical or nearly identical to those of gene B (48, 143, and 1,001 bp, respectively). This is in agreement with the observation that POMC mRNAs A and B are detected as a single band on northern blots (Martens et al. 1985). Since intron B of gene A (2,862 bp) is only slightly larger than intron B of gene B (2,403 bp), the difference in length between the two
Xenopus POMC genes is mainly caused by different lengths of introns A (4,367 bp in gene A and 2,435 bp in gene B).

Internal Repeats and Repetitive Elements in Xenopus POMC Gene A

In Xenopus POMC gene A three stretches were found that contain so-called simple DNA sequences. The first region (nucleotides 1234–1464) contains a contiguous stretch of 35 ATs and 22 TTTCs, the second stretch (nucleotides 4766–4813) is AT-rich, while the third region (nucleotides 8949–9254) contains 60 times the tetranucleotide TAGA (regions I, VIII, and XII, respectively, in fig. 2). Repetitive TAGA sequences have been identified both in rat repetitive DNA and in sex-specific satellites.
of *Drosophila* and mouse (Alonso et al. 1983; Singh et al. 1984). Stretches of AT dinucleotides have been found in a number of organisms, such as yeast, *Drosophila*, and trypanosoma (Delange et al. 1984; Lorincs and Reed 1984; Tautz and Renz 1984), and they have been shown to occur as interspersed repeats in the *Xenopus* genome (Greaves and Patient 1985). Alternating pyrimidines and purines, such as the \((\text{AT})_3\) stretch in the POMC gene, can adopt a left-handed (Z) DNA configuration (Wells et al. 1982), and as such they might influence gene transcription (Naylor and Clark 1990).

Besides these simple sequences, we found in POMC gene A two different repetitive elements—namely, members of the previously described families of JH12 repetitive elements (Meyerhof et al. 1987) and vitellogenin (Vi) transposon–like elements (Schubiger et al. 1985). Intron A contains two JH12 elements (nucleotides 1981–2006 and 4100–4351; regions III and VII, respectively, in fig. 2), of which one is inverted relative to the other. In addition, intron A contains three Vi elements (nucleotides 1981–2006, 4100–4351, and 6300–6351; regions III and VII, respectively, in fig. 2).
FIG. 2.—Schematic representation of locations of repeats in *Xenopus* POMC gene A and its 5'- and 3'-flanking regions. The repeats are indicated by roman numbers: I, VIII, and XII are simple sequences; II, IV, VI, and X are Vi-transposon-like elements; III and VII are JH 12-repetitive elements; IX is a cluster of internal repeats; V is an inverted repeat of a part of IX; and XI is an internal tandem repeat. For reference, the locations of the three exons of the *Xenopus* POMC gene are indicated by closed boxes.

2615–3014, and 3816–4100; regions II, IV, and VI, respectively, in fig. 2), while a fourth Vi element is found in intron B (nucleotides 6396–6857; region X in fig. 2). The most upstream Vi element is inverted relative to the other three elements. Most of the length difference between introns A (1,932 bp) of POMC genes A and B can be attributed to the occurrence of the three Vi-repetitive elements (1,173 bp) in intron A of gene A, because gene B does not contain a Vi element.

Analysis of *Xenopus* POMC gene A for internal repeats, by comparison of the gene with itself, revealed that, besides the repetitive elements mentioned above, intron B contains a cluster of four repeats (nucleotides 5908–6149, 6150–6395, 6858–7110, and 7111–7320; region IX in fig. 2) just upstream of exon 3; one of the four Vi transposon-like elements is present between the second and third repeat. Part of the cluster (nucleotides 6146–6205) is inversely repeated in intron A (nucleotides 3610–3676; region V in fig. 2). Furthermore, just downstream of exon 3 a 24-bp sequence is four times tandemly repeated (nucleotides 8535–8652; region XI in fig. 2). Neither the clustered repeats nor the tandemly repeated sequence shows significant identity with any sequence in the EMBL or GenBank data bases.

To determine the relative occurrence of the described repetitive elements in the *Xenopus* genome, we hybridized Southern blots of DNA fragments of *Xenopus* POMC gene A with a probe derived from *Xenopus* genomic DNA. This analysis (data not shown) confirmed that the JH 12 and Vi repetitive elements belong to families that are highly and moderately repetitive, respectively (Schubiger et al. 1985; Meyerhof et al. 1987). In this analysis no other regions in gene A gave a hybridization signal, including the clustered repeats upstream of exon 3.

Comparative Structural Analysis of *Xenopus* POMC Genes A and B

To find conserved and thus potentially important regulatory sequences, we examined dot matrices of *Xenopus* POMC genes A and B. This analysis revealed, besides the exons, three regions of high similarity. The promoter regions are >90% identical. We have previously identified putative cis-acting regulatory sequences conserved among the promoters of *Xenopus* POMC gene B and of four mammalian POMC genes (Deen et al. 1991) whereby sequences of four or more identical nucleotides and >65% nucleotide sequence identity among all species were classified as conserved regions. By the same criteria, it appears that all but one (box A) of these sequences are also present in *Xenopus* POMC gene A. This finding further suggests that these sequences are involved in the regulation of POMC gene transcription.

The second region of high similarity between the *Xenopus* POMC genes concerns JH12 repetitive elements. POMC gene B contains a JH12 element (Deen et al. 1991) at approximately the same position as the most upstream JH12 element in gene A (region III in fig. 2). In the *Xenopus* genome, JH12 repetitive elements can occur
Proopiomelanocortin Genes A and B of *Xenopus laevis* 489

dispersed as well as clustered (Meyerhof et al. 1987; Deen et al. 1991). These elements might be involved in the control of gene activation, because JH12 transcripts are only detected after a specific developmental stage (midblastula transition; Meyerhof et al. 1987). The alignment of the two JH12 repetitive elements of POMC gene A with both the element of gene B and the JH12 consensus sequence compiled from nine JH12 copies present in eight *Xenopus* genes (Deen et al. 1991) is shown in figure 3. Both repeats in POMC gene A are >86% identical with the JH12 consensus sequence. Figure 3 also reveals that the most downstream JH12 element of gene A (nucleotides 4100–4351; region VII in fig. 2) is interrupted by one of the four Vi-repetitive elements (region VI in fig. 2). In addition, as compared with the JH12 consensus sequence, this JH12 element lacks 54 bp exactly at the position corresponding to the apparent insertion site of the Vi-repetitive element. These observations suggest that during evolution the insertion of this Vi-repetitive element in POMC gene A occurred after insertion of the JH12 repetitive element and that this insertion induced a 54-bp deletion within the JH12 element. Two of the four Vi elements invaded in repeats—namely, in the repeats upstream of exon 3 and in the second JH12 element.

The third region of high similarity concerns the four clustered repeats upstream of the third exon of POMC gene A (nucleotides 5908–6149, 6150–6395, 6858–7110, and 7111–7320, referred to as C\(^{*}\), C\(^{2*}\), C\(^{4*}\) and D*, respectively; region IX in fig. 2) that are 63%–72% identical with a 500-bp region (VI) upstream of exon 3 in POMC gene B. The identity between the clustered repeats and region VI is relatively low because of several internal deletions/insertions. The flanking regions of the clustered repeats and of region VI (the 54 bp upstream of the clustered repeats and of region VI and the 58 bp downstream of them) are also conserved. On the basis of our dot matrix analysis and the search for internal repeats in gene A, it is likely that before genome duplication in *Xenopus* a segment-Z element was present in the POMC gene (fig. 4). This element probably consisted of a repeat-C and a repeat-D sequence. After genome duplication a tripling of the repeat-C sequence occurred in POMC gene A, giving rise to C\(^{*}\), C\(^{2*}\), and C\(^{4*}\). Next, an invasion of a Vi-repetitive element must have occurred between repeats C\(^{*}\) and C\(^{4*}\) of POMC gene A. Alternatively, deletion events in POMC gene B can explain the observed relationship between these intron regions. Besides the three large conserved sequences mentioned above, other (short) conserved regions are found. With respect to the exons and the large conserved regions mentioned above, their relative positions are similar in both genes. Finally, dot matrix analysis of the POMC gene A with the inverse complement of the POMC gene B revealed significant similarity only between JH12 elements and simple sequences (data not shown).

Our analysis shows that *Xenopus* POMC genes A and B occur as a closely related pair, as do the *Xenopus* genes encoding Vi’s (Wahli et al. 1979), adult globins (Widmer et al. 1981), albumins (Westley et al. 1981), proenkephalins (Martens and Herbert 1984), and L1 ribosomal proteins (Loreni et al. 1985). Each pair displays 4%–9% divergence in the corresponding exon sequences. This supports the hypothesis that ~30 Mya the ancestor of *X. laevis* underwent a genome duplication (Bisbee et al. 1977; Thiebaud and Fischberg 1977). Although the corresponding exon sequences of these gene pairs are usually similarly sized, the lengths of the corresponding intron sequences can vary significantly. This is also the case in the *Xenopus* POMC genes, in which the corresponding exons and introns B are about equally sized but in which introns A differ considerably. As already mentioned, this length difference between
Fig. 3.—Alignment between nucleotide sequences of JH12-repetitive elements in *Xenopus* POMC genes A and B [lines (1)–(3)] and JH12 consensus sequence [line (4)]. A colon indicates a gap that has been introduced to increase similarity. An asterisk (*) indicates identity between all four JH12 sequences. The terminal inverted repeat of the JH12 consensus sequence is underlined by arrows. A nucleotide below the consensus sequence represents an equally well-conserved nucleotide. Vi-III indicates the position of the third Vi transposon-like element of POMC gene A. The JH12 nucleotide sequences of *Xenopus* POMC gene B and the JH12 consensus sequence have been taken from the study by Deen et al. (1991).
FIG. 4.—Possible evolutionary relationship between regions upstream of third exon of POMC genes A and B. An asterisk (*) indicates an evolutionary relationship with the ancestral POMC gene. For reference, the regions upstream of the third exon of POMC genes A and B are indicated by roman numbers, according to the protocols of fig. 2 and Deen et al. (1991), respectively.

The two A introns can primarily be attributed to insertion/deletion events of repetitive Vi elements. Such events may also account for the variation in corresponding intron sequences of the *Xenopus* albumin and Vi gene loci (Ryffel et al. 1983). Repetitive elements located at approximately the same positions in a pair of genes, such as the JH12 element in the POMC genes, may already have been present in the ancestral gene. If these elements have a function in the regulation of gene expression, one would expect that they do not diverge as fast as other parts of intron sequences and thus that these elements will increase the similarity between corresponding intron sequences. It is remarkable that the corresponding intron sequences of paired genes coding for Vi’s B (Germond et al. 1983) and albumins (May et al. 1983) are very similar, while there is no significant similarity between introns of the paired genes encoding Vi(s) A or larval β-globins (Germond et al. 1983; Meyerhof et al. 1986). The *Xenopus* POMC genes contain many conserved regions in their introns, and thus they seem to belong to the group of relatively highly conserved genes that includes the Vi(s) B and albumin genes. At present, however, the functional significance of the conserved intron se-
quences is unclear. The analyses with Vi B and albumin genes were performed with heteroduplex and R-loop mapping techniques, and thus it is not known whether the high similarity between the corresponding introns of these genes can be attributed to the presence of conserved repetitive elements. Since, with the notable exception of the JH12 element, the repetitive elements are not conserved in the *Xenopus* POMC genes, these elements might not have a function in the regulation of POMC gene transcription.

In conclusion, our screening of a *X. laevis* genomic library resulted in the isolation and characterization of a second POMC gene, gene A. POMC genes A and B contain, in addition to their exons, three large conserved regions: the promoter region, a JH12 repetitive element, and a region just upstream of exon 3. It is interesting to note that POMC genes A and B are coordinately transcribed to approximately the same level in one cell type—namely, the melanotrope cell of the intermediate lobe of the toad (Martens 1986). It may well be that the three conserved regions are involved in this coregulated expression of the two POMC genes. We are currently studying by means of a homologous gene transfer system—namely, microinjection of the *Xenopus* POMC gene and mutants thereof into *Xenopus* oocytes and eggs—possible functions of the conserved elements in POMC gene regulation.

**Acknowledgments**

This research was supported by National Institutes of Health grant R01-DK-37902, by the Netherlands Organization for Scientific Research (NWO), and by European Community grant ST25-0468-C. We want to thank the CAOS/CAMM personnel for computer assistance.

**LITERATURE CITED**


JAN KLEIN, reviewing editor

Received June 28, 1991; revision received September 9, 1991

Accepted September 9, 1991