

SpCOUP-TF: A sea urchin member of the steroid/thyroid hormone receptor family

SIU-MING CHAN, NAIDI XU, CHRISTINA C. NIEMEYER, JAMES R. BONE, AND CONSTANTIN N. FLYTZANIS*

Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030

Communicated by Roy J. Britten, July 14, 1992 (received for review March 25, 1992)

ABSTRACT Embryonic nuclear proteins from the sea urchin *Strongylocentrotus purpuratus* bind *in vitro* to a cis-acting element that lies upstream of the actin gene *CyIIIb* and consists of two direct repeats homologous to steroid hormone response elements. This sea urchin element is specifically recognized by transcription factor COUP-TF from HeLa cell nuclear extracts as well. A sea urchin gene homologous to the human COUP-TF1 gene was detected by blot hybridization to *S. purpuratus* genomic DNA. Screening of a genomic DNA library with a human COUP-TF1 cDNA probe produced overlapping genomic clones that carry the *S. purpuratus* gene. Our results indicate that the sea urchin homologue has a structure similar to the human COUP-TF1 gene, i.e., conserved intron positions characteristic of this subgroup of steroid hormone receptors. Sequencing of the exons that encode the DNA- and ligand-binding domains of the sea urchin protein revealed 96% and 92% amino acid identity to the domains of the human protein, respectively. Transcripts derived from the *S. purpuratus* COUP-TF homologous gene were detected in ovarian and embryonic RNAs from various stages.

The protein COUP-TF (chicken ovalbumin upstream promoter-transcription factor) is required for expression of the chicken ovalbumin gene (1, 2), binds specifically to the rat insulin II gene promoter (3), and belongs to the steroid/thyroid hormone receptor superfamily as determined by its structure and the extensive sequence homology of its DNA- and ligand-binding domains (4, 5). A specific ligand for COUP-TF has not yet been identified, but its activation in transfection assays was achieved by physiological concentrations of dopamine, suggesting an activation of COUP-TF through this membrane receptor pathway (6). The *Drosophila* seven-up (*svp*) gene was shown to be the fly homologue of the human COUP-TF1 gene, and loss of *svp* function causes an apparent change of cell fate of certain photoreceptor cell precursors during eye development (7). Another member of the human COUP-TF subfamily (ARP-1 or COUP-TF2) has been isolated (8, 9) and shows an 87% overall amino acid identity to COUP-TF1. ARP-1 was shown to down-regulate the apolipoprotein AI gene and bind to the regulatory regions of the chicken ovalbumin gene, rat insulin II gene, and human apolipoprotein B and CIII genes (8). COUP-TF1 was shown to bind specifically to a COUP-like element which is part of a negative regulatory region in the long terminal repeat of human immunodeficiency virus type 1 (10). Thus, it appears that the COUP-TFs are capable of both up- and down-regulating the activity of a variety of genes in a plethora of tissues.

The *Strongylocentrotus purpuratus* actin gene *CyIIIb* is activated during the cleavage stages of the sea urchin embryo (11), in the cells destined to become the aboral ectoderm (12). The upstream region of *CyIIIb* is necessary for proper expression in microinjected embryos and contains at least

five different elements (13). Within its regulatory region (position -440), we identified a COUP-TF binding site which we call C1R. This site is part of a more complex element, C1, as shown by footprinting and gel mobility-shift assays (13). Deletion of the C1 element results in a 2-fold reduction of activity from the *CyIIIb* promoter in microinjected embryos, suggesting a possible *in vivo* role of this element in the regulation of the gene (13). Intrigued by the extensive sequence conservation between the human COUP-TF1 gene and the fruit fly *svp* gene, we used the human COUP-TF1 cDNA as a probe to isolate and characterize the homologous gene from *S. purpuratus*.[†] Our data show that this gene encodes a sea urchin member of the steroid/thyroid receptor superfamily (14).

METHODS

Preparation of Nuclear Extracts. *S. purpuratus* embryos were cultured at $1-2 \times 10^7$ embryos per liter of artificial seawater in stirring vessels with constant aeration at 15°C. Nuclei and nuclear protein extracts were prepared as described (15, 16).

Genomic DNA Blots. High molecular weight genomic DNA was prepared from sperm of individual *S. purpuratus* (17). Digated DNA (10 µg) was electrophoresed in 1% agarose gels and transferred onto nitrocellulose membranes (18). The membranes were prehybridized in 0.5 M sodium phosphate buffer (pH 6.8) with 1% SDS, 2 mM EDTA, 1× Denhardt's solution (19), and calf thymus DNA (100 µg/ml) for 2 hr at 60°C. Hybridization took place in the same buffer containing denatured probe (10⁶ cpm/ml) at 60°C for 16–18 hr. The membranes were washed in 0.1 M sodium phosphate buffer, pH 6.8/0.2% SDS/2 mM EDTA twice (20 min each) at room temperature and twice at 50°C. The final two washes (30 min each) were in 0.05 M sodium phosphate buffer, pH 6.8/0.1% SDS/1 mM EDTA at 60°C.

Mobility-Shift Assays. The reaction mixture (10 µl) contained 10 µg of nuclear proteins, 5000–10,000 cpm of ³²P-labeled oligonucleotide probe, 2.5 µg of poly(dA·dT) as nonspecific competitor, 1 mM MgCl₂, 50 mM KCl, 20 mM Hepes (pH 7.8), 5 mM dithiothreitol, and 10% (vol/vol) glycerol. The samples were incubated on ice for 15 min and electrophoresed in 6% polyacrylamide gels in Tris borate/EDTA buffer (20) at 4°C.

Genomic Library Screening. A *S. purpuratus* genomic library constructed with DNA from individual "B" (21) in the vector EMBL3 was screened with the 1.5-kilobase (kb) human COUP-TF cDNA insert (5) as described (22). The prehybridization, hybridization, and washing conditions were the same as described for the genomic DNA blots.

Sequencing. Genomic fragments from clones λ3 and λ4 were subcloned in the vector pBluescript (Stratagene) and their sequences were determined by the chain-terminating

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*To whom reprint requests should be addressed.

[†]The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L01104 and L01105).

dideoxynucleotide method (23) with the modifications adopted for double-stranded plasmid templates (24).

Reverse Transcription-PCR. Total RNA was extracted from embryos and ovarian tissue (25). Reverse transcription was performed with 10 μ g of total RNA and 0.1 μ g of (dT)₂₀ in 50 μ l containing 100 mM Tris (pH 8.3), 10 mM MgCl₂, 50 mM KCl, 14 mM 2-mercaptoethanol, 5 units of RNAGuard (Pharmacia), and 20 units of avian myeloblastosis virus reverse transcriptase (Pharmacia). The reaction mixture was incubated at 42°C for 2 hr, followed by inactivation at 95°C for 5 min. The PCR mix (50 μ l) contained 2.5 μ l of the cDNA reaction mix, 2 mM MgCl₂, 10 mM Tris (pH 8.3), 50 mM KCl, 1 mM dNTPs, 1 μ g of each primer, and 5 units of *Taq* DNA polymerase. The PCR conditions were 1 min at 95°C, 1 min at 60°C, and 2 min at 72°C. Ten microliters of the PCR mix was loaded on a 1.5% agarose or 6% polyacrylamide gel and the DNA products were analyzed by blot hybridization.

RESULTS

The *CyIIIb* Upstream Element C1 Contains a COUP-TF Binding Site. The sequence homology between the C1 element (AGGTCAGCGGGTCA, Fig. 1A) and the COUP element of the chicken ovalbumin gene (GTGTCAAAGGTCA), as well as the spacing between the AGGTCA half-sites (26), suggested that COUP-TF may bind to the C1 element or, conversely, the protein(s) that binds to C1 may also bind to the COUP element. To explore these possibilities, we performed mobility-shift assays using either *S. purpuratus* embryonic nuclear proteins, which were previously shown to protect the C1 site (13), or human HeLa cell nuclear proteins. With the C1 oligonucleotide as probe, several specific protein-DNA complexes were formed (Fig. 1B Left). The mutant oligonucleotide C1mut at 40-fold excess partially competed with the C1 probe for protein binding, suggesting that

the introduced substitutions were not sufficient to abolish binding and thus, that some of these positions (Fig. 1A) are important for recognition. When 100-fold excess COUP competitor was added, a number of complexes still remained, suggesting that in addition to COUP-TF (bands marked with an asterisk in Fig. 1B), other nuclear proteins recognize the C1 element. When we used HeLa cell nuclear extract, a number of specific complexes were also formed. Competition with the COUP element of the ovalbumin gene eliminated the COUP-TF-specific complexes, but only the unlabeled C1 oligonucleotide successfully competed with all the bands. Thus, HeLa nuclear extract also contains proteins other than COUP-TF which specifically bind the C1 element (presumably the C1L palindrome). We repeated the previous experiment by replacing the C1 probe with the COUP probe. Both the sea urchin and HeLa cell extracts formed specific COUP-TF complexes only (data not shown).

The C1 element was divided into two smaller fragments by synthesizing two new oligomers (C1R and C1L). C1R consists of 22 nucleotides from the 3' end of C1 containing the AGGTCA repeats, whereas C1L consists of 20 nucleotides from the 5' end including the imperfect palindrome TTTG-GCGGAAA. Mobility-shift assays with either C1L or C1R (Fig. 1B Center) indicated that the corresponding binding proteins are distinct, since the two oligomers did not compete. Competition with unlabeled COUP was partial at 40-fold excess, suggesting that sea urchin binding factor has greater affinity for the C1R element than for the ovalbumin COUP element. To further characterize the binding specificity of the C1R site, we made another mutant oligonucleotide (mutlr) by replacing the two cytosines at positions -433 and -441 with guanines. The result of these two substitutions was complete abolishment of binding (Fig. 1B Right), in agreement with the significant contact sites of human COUP-TF on the ovalbumin promoter (3). These results

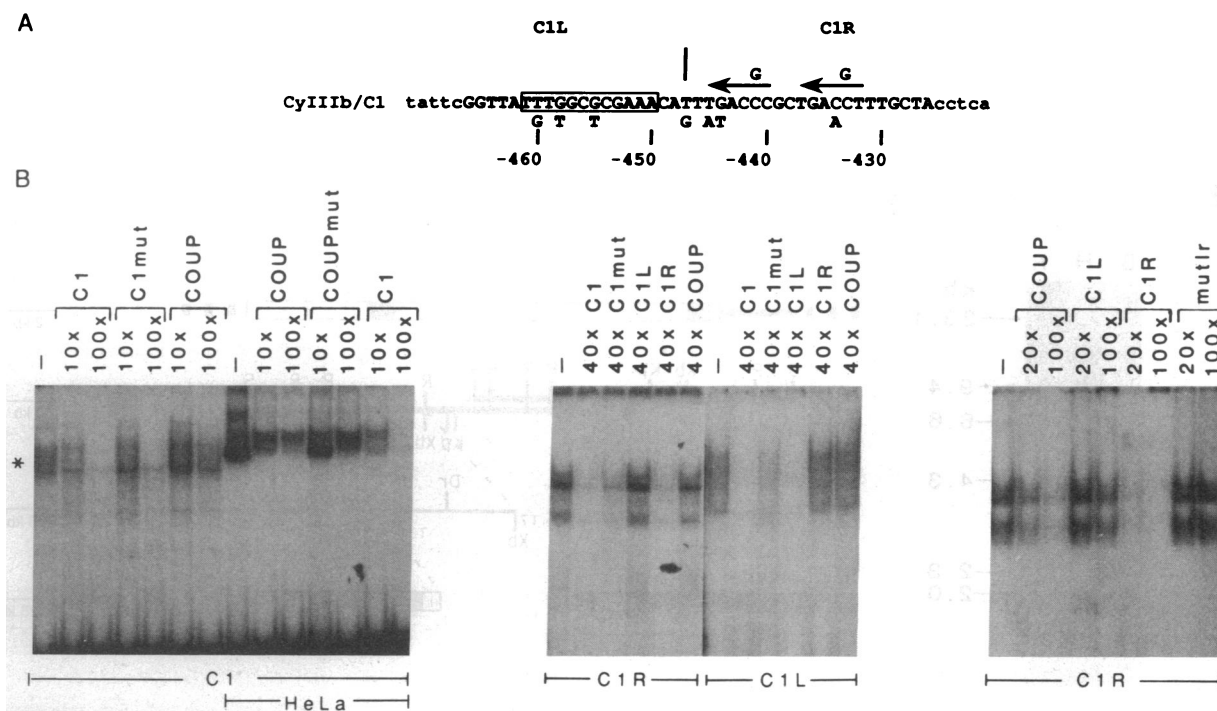


FIG. 1. COUP-TF specific binding upstream of *CyIIIb*. (A) The C1 sequence upstream of *CyIIIb* from -421 to -471 is presented. Uppercase letters correspond to the "footprinted" region (13). The imperfect palindrome of C1L is boxed, and arrows depict the direct repeats of the steroid response element of the C1R site. The vertical bar separates the two elements used in B. The seven nucleotide substitutions of the C1mut oligonucleotide are presented below the sequence, whereas the two substitutions in the C1R element are shown above the two arrows. (B) Mobility-shift assays with the C1 element. The probe used in each assay is denoted below the autoradiogram. All lanes contained *S. purpuratus* 24-hr embryonic nuclear extracts except for the ones denoted HeLa. Numbers above each lane correspond to the fold excess of unlabeled oligonucleotide competitor used in each case; -, no competitor. Asterisk marks the COUP-TF-specific bands.

constitute strong evidence that the sea urchin protein which binds the C1R element is homologous to human COUP-TF.

COUP-TF Homologous Genes Are Present in the *S. purpuratus* Genome. Sea urchin genomic DNA blots with the human COUP-TF1 cDNA as probe (Fig. 2A) revealed strong and weak hybridization signals, which could be either relative to the degree of homology of the probe to different genes and/or relative to the size of the hybridizing fragments. The number of observed bands should not correspond to the number of genes, because the probe spans almost the entire human cDNA and thus hybridizes to genomic fragments belonging to exons separated by introns. A 4.3-kb *EcoRI* band, which gives a strong hybridization signal, carries the exon that encodes the SpCOUP-TF hormone-binding domain and shows the highest degree of homology. Other bands must represent another, less homologous gene(s) and the DNA-binding domain of SpCOUP-TF. The relatively small number of bands indicates the presence of only a few homologous genes.

Using the human COUP-TF cDNA probe, we screened about 300,000 plaques of a sea urchin genomic library and obtained 6 positive clones (λ 1– λ 6) that showed different hybridization intensity signals. Detailed restriction mapping of λ 3 and λ 4 indicated that these two clones were partially overlapping and carried sequences homologous to the human COUP-TF cDNA on their nonoverlapping regions (Fig. 2B). The sequencing results confirmed that the DNA-binding domain was encoded on λ 3 and the ligand-binding domain on λ 4. These two domains are separated by the first intron of the gene, which is found at the same position (Arg-Arg-Glu . . . intron 1 . . . Ala-Val-Gln) as in human COUP-TF1 (27). The second intron, which divides the ligand-binding domain, is found also at a conserved position. The sequence of the subclones, which includes the entire first and second exons as well as part of the first and second introns, is presented together with the predicted amino acid sequence in Fig. 3 A and B. We are missing the third exon of the SpCOUP-TF gene and thus the carboxyl terminus of the protein. The two exons encode 381 amino acids, or about 81% of the total protein, if the length of the carboxyl terminus is as conserved in the sea urchin as it is between the human COUP-TF and the *Drosophila svp* gene. The overall identity of the protein sequence

at hand to the human COUP-TFs (hCOUP-TF1 and hCOUP-TF2 or ARP-1) is 73% (Fig. 3C). Within the DNA-binding domain (the 66 amino acids which include the two zinc fingers) and the ligand-binding domain (the rest of the carboxyl-terminal part of the protein) the identity of the amino acid sequence is 96% and 92%, respectively. This impressive structural identity indicates a functional conservation in addition to the recognition of the same element. The extensive similarity of the ligand-binding domains suggests that these molecules are activated by the same hormone or the same pathway. The amino acid identity of SpCOUP-TF to the *Drosophila svp* gene product (7) is also 96% for the DNA-binding and 92% for the ligand-binding domain. A possible functional conservation of the COUP-TFs that extends from protostomes to such diverse deuterostomes as sea urchins and humans may indicate a common signaling pathway. Especially interesting would be the role of this molecule during early embryonic development in cell fate determination, if a generalization of the *svp* function applies to other animals.

SpCOUP-TF Is Expressed During Early Embryonic Development. RNA blot hybridizations using total RNA isolated from ovaries and various embryonic stages were suggestive of the presence of SpCOUP-TF transcripts but inconclusive, most probably because of the low abundance of the transcripts. We thus employed a reverse transcription-PCR method to detect the SpCOUP-TF transcripts. Fig. 4B shows our strategy in selecting appropriate primers for the PCR. To avoid amplification of contaminating genomic DNA in our RNA preparations, the two PCR primers (Z and Dex), which anneal to the DNA- and ligand-binding domains, respectively, are separated by about 13 kb of intron sequence, so that even if genomic DNA were amplified the products would not interfere with the band obtained from the mRNA, which is 453 bases long. The use of bacterial rRNA, omission of reverse transcriptase from the reaction, or employment of a different primer in this assay did not result in any synthesis, indicating the specificity of the reaction. A product of the expected size was detected (Fig. 4A) by blot hybridization of the PCR products with an internal fragment of the ligand-binding domain as a probe (Fig. 4B). When different sets of specific primers (S and Dex or A and Dex) were used in the

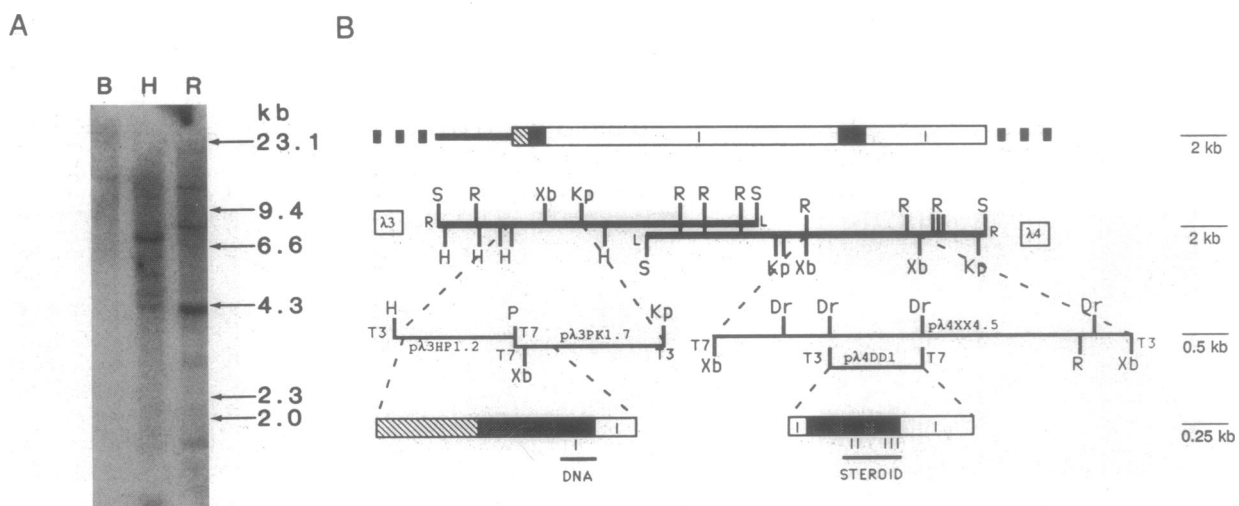


FIG. 2. Characterization of the *S. purpuratus* COUP-TF (SpCOUP-TF) gene. (A) Restriction enzyme-digested (B, *Bam*HI; H, *Hind*III; R, *Eco*RI) genomic DNA from individual "Buster" hybridized with the human COUP-TF cDNA probe. Molecular size standards were obtained by digesting wild-type λ phage DNA with *Hind*III. (B) Restriction enzyme maps of the overlapping clones λ 3 and λ 4 and the subclones used for sequencing. In a schematic representation of the gene, the coding sequences are presented as gray and black boxes, the presumed 5' untranslated region as a hatched box, and the upstream region as a solid line. Introns are depicted as open boxes marked "I" and the conserved domains of the protein as black boxes (I, DNA-binding domain; II and III, steroid-binding domains). R or L on each side of the genomic clones corresponds to the arms of the bacteriophage, and T3 or T7 denotes the orientation of the inserts within the polylinker of the subcloning vector pBluescript. Dr, *Dra* I; H, *Hind*III; Kp, *Kpn* I; P, *Pst* I; R, *Eco*RI; S, *Sal* I; Xb, *Xba* I.

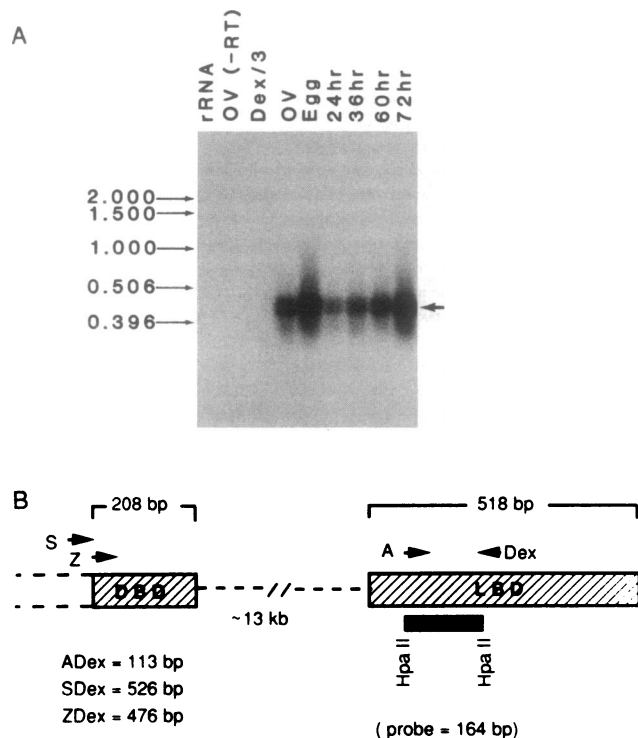


FIG. 4. Expression of the SpCOUP-TF gene during embryonic development. (A) DNA blot hybridization to reverse transcription-PCR products obtained from RNAs isolated from ovaries (OV), eggs, and embryos (24, 36, 60, and 72 hr postfertilization). The cDNAs were amplified with the oligonucleotides Z and Dex (see B), and the *Hpa* II fragment, which does not overlap with the primers, was used as probe. The expected band is marked by an arrow. rRNA, bacterial ribosomal RNA; OV (-RT), the template was ovarian RNA but the reverse transcriptase was omitted from the reaction; Dex/3, the primer Dex was used in combination with the non-COUP-TF-specific oligonucleotide 3 and ovarian RNA as template. (B) Positions of the PCR primers, the expected size of amplified products, and position of the *Hpa* II fragment used as probe. Only the products of the Z/Dex combination are presented in A, although the A/Dex and S/Dex products have also been obtained. DBD, DNA-binding domain; LBD, ligand-binding domain.

The SpCOUP-TF gene is expressed at very low levels, during oogenesis and embryogenesis, perhaps increasing by the pluteus stage. In preliminary experiments using somatic tissues, other than ovary, SpCOUP-TF transcripts were also detected in the adult animal. Although we have no clue as to what other target genes this receptor may be regulating, it is possible that SpCOUP-TF is not exclusively participating in an embryonic regulatory network. In this respect the function of this transcription factor/hormone receptor may be more generalized than that of its *Drosophila* counterpart, the *svp* gene product. Is then its mode of activation ligand-dependent throughout the life cycle of the sea urchin? Very little is known about the presence of steroid hormones and steroid-metabolizing enzymes in sea urchins (32). Estradiol and progesterone were found in the ovaries of the sea urchin *Strongylocentrotus franciscanus* (33), and estradiol was shown to affect cleavage, nucleic acid metabolism, and protein synthesis in *S. purpuratus* embryos (34). If SpCOUP-TF is a receptor for a steroid hormone, identification of its activating ligand would be of great importance in elucidating the function of signaling pathways during sea urchin embryogenesis.

We are indebted to the laboratories of Drs. Ming-Jer Tsai and Bert W. O'Malley for providing us with the human COUP-TF cDNA clone. We gratefully acknowledge the critical review of this manuscript by Drs. Ming-Jer Tsai and Yassemi Capetanaki. This work was supported by National Institutes of Health Grant HD22055 to C.N.F.

- Sagami, I., Tsai, S. Y., Wang, H., Tsai, M.-J. & O'Malley, B. W. (1986) *Mol. Cell. Biol.* **6**, 4259-4267.
- Bagchi, M. K., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. (1987) *Mol. Cell. Biol.* **7**, 4151-4158.
- Hwang, Y.-P., Wang, L.-H., Tsai, S. Y. & Tsai, M.-J. (1988) *J. Biol. Chem.* **263**, 13470-13474.
- Miyajima, N., Kadowaki, Y., Fukushige, S., Shimizu, S., Semba, K., Yamanashi, Y., Matsubara, K., Toyoshima, K. & Yamamoto, T. (1988) *Nucleic Acids Res.* **16**, 11057-11074.
- Wang, L.-H., Tsai, S. Y., Cook, R. G., Beattie, W. G., Tsai, M.-J. & O'Malley, B. W. (1989) *Nature (London)* **340**, 163-166.
- Power, R. F., Lydon, J. P., Conneely, O. M. & O'Malley, B. W. (1991) *Science* **252**, 1546-1548.
- Ulodzik, M., Hiromi, Y., Weber, U., Goodman, C. S. & Rubin, S. U. (1990) *Cell* **60**, 211-224.
- Ladiaz, J. A. A. & Karathanasis, S. K. (1991) *Science* **251**, 561-565.
- Wang, L.-H., Ing, N. H., Tsai, S. Y., O'Malley, B. W. & Tsai, M.-J. (1991) *Gene Exp.* **1**, 207-216.
- Cooney, A. J., Tsai, S. Y., O'Malley, B. W. & Tsai, M.-J. (1991) *J. Virol.* **65**, 2853-2860.
- Flytzanis, C. N., Bogosian, E. A. & Niemyer, C. C. (1989) *Mol. Reprod. Dev.* **1**, 208-218.
- Cox, K. H., Angerer, L. M., Lee, J. J., Davidson, E. H. & Angerer, R. C. (1986) *J. Mol. Biol.* **188**, 159-172.
- Niemeyer, C. C. (1992) Doctorate thesis (Baylor College of Medicine, Houston, TX).
- Evans, R. M. (1988) *Science* **240**, 889-895.
- Morris, G. F. & Marzluff, W. F. (1983) *Biochemistry* **22**, 645-653.
- Morris, G. F., Price, D. H. & Marzluff, W. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3674-3678.
- Lee, J. J., Shott, R. J., Rose, S. J., II, Thomas, T. L., Britten, R. J. & Davidson, E. H. (1984) *J. Mol. Biol.* **172**, 149-176.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517.
- Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun.* **23**, 641-646.
- Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560-564.
- Calzone, F. J., Lee, J. J., Le, N., Britten, R. J. & Davidson, E. H. (1988) *Genes Dev.* **2**, 305-318.
- Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180-182.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165-170.
- Posakony, J. W., Flytzanis, C. N., Britten, R. J. & Davidson, E. H. (1983) *J. Mol. Biol.* **167**, 361-389.
- Cooney, A. J., Tsai, S. Y., O'Malley, B. W. & Tsai, M.-J. (1992) *Mol. Cell. Biol.* **12**, 4153-4163.
- Ritchie, H. H., Wang, L.-H., Tsai, S., O'Malley, B. W. & Tsai, M.-J. (1990) *Nucleic Acids Res.* **18**, 6857-6862.
- Umesono, J., Murakami, K. K., Thompson, C. C. & Evans, R. M. (1991) *Cell* **65**, 1255-1266.
- Yu, V. C., Delsert, C., Anderson, B., Holloway, J. M., Devary, O. V., Näär, A. M., Kim, S. Y., Boutin, J.-M., Glass, C. K. & Rosenfeld, M. G. (1991) *Cell* **67**, 1251-1266.
- Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J.-Y., Staub, A., Garnier, J.-M., Mader, S. & Chambon, P. (1992) *Cell* **68**, 377-395.
- Segraves, W. A. (1991) *Cell* **67**, 225-228.
- Lehoux, J.-G. & Sandor, T. (1970) *Steroids* **16**, 141-171.
- Botticelli, C. R., Hisaw, F. L., Jr., & Wotiz, H. H. (1961) *Proc. Soc. Exp. Biol. Med.* **106**, 887-889.
- Jolley, W. B., Martin, W. E., Bamberger, J. W. & Stearns, L. W. (1962) *J. Endocrinol.* **25**, 183-188.