How I became one of the fathers of a superfamily

Pierre Chambon

Elucidating the mechanisms underlying transcriptional control in vertebrates has been my main research project for almost 45 years. It originated from my M.D. thesis work (1958) on the effects of X-ray irradiation on free ribonucleotides in rat tissues, which suggested that RNA could be synthesized from nucleoside triphosphates. A year later, the field of *in vitro* transcription opened up with the discovery of a nucleoside triphosphate–dependent RNA polymerase activity in rat liver nuclei[1]. However, the mechanism of DNA-directed RNA synthesis was established in the 1960s using purified bacterial RNA polymerase, because only little soluble activity could be recovered from vertebrate nuclei; the bulk of RNA polymerase, the 'aggregate' enzyme, was transcriptionally engaged and tightly bound to chromatin[2]. Nevertheless, two important observations emerged from my studies with such nuclear preparations. First, in 1963, I discovered poly(ADP-ribose) polymerase[3], which has recently been shown to belong to a multigene family with highly diverse functions[4,5]. Second, and more decisively for my future, I found that administration of estradiol to immature chickens elicited an increase of liver 'aggregate' polymerase, which preceded the induction of protein synthesis[6]. This finding suggested that the bound estrogen receptor (ER), just discovered by Elwood Jensen, could act positively on the initiation of transcription.

As interesting as this hormonal induction was, there was no hope of explaining its molecular mechanism in the foreseeable future. The complexity of vertebrate genomes and the lack of molecular approaches precluded the prerequisite characterization and isolation of promoter regions of the responding genes. Why characterize RNA polymerase in animals if it could not be used to study how the expression of their genetic programs is controlled? Furthermore, the opinion at the time was that the mechanisms involved would be simple variations upon those already established for prokaryotes. Was it therefore wise to pursue such an ill-fated project?

Nevertheless, the spatiotemporal regulation of transcription in multicellular organisms was clearly a huge problem that is not encountered in prokaryotes. This was particularly acute during embryonic development, when many cells with diverse functions derive from a single cell. The expression of each gene could not be controlled by a specific regulatory protein, because controlling its expression would require another regulatory protein, the expression of which would in its turn require another one and yet another one. The control of transcription in animals had therefore to be combinatorial in nature, involving regulatory proteins that, on their own, would rarely be specific for a given gene.

Because this mechanism would probably require an RNA polymerase more sophisticated than the bacterial one, I decided that studying RNA polymerase in vertebrates was worthwhile, particularly given that Stirpe and Fiume had shown in 1967 that the toxin α-amanitin selectively inhibited only a fraction of RNA synthesis in isolated mouse liver nuclei[7,8]. With four graduate students—Claude Kédinger, Jean-Louis Mandel, Francis Gissinger and Madeleine Cochet-Meilhac—I resumed in 1968 my attempts to characterize the animal RNA polymerase.

Multiplicity of eukaryotic RNA polymerases

In 1969, Roeder and Rutter[9] and our group[10] independently succeeded in solubilizing the 'aggregate' enzyme and discovered several RNA polymerase (Pol) activities. We isolated two activities from calf thymus, Pol A and Pol B, of which only Pol B was inhibited by α-amanitin[10], whereas Roeder and Rutter[9] separated Pol I, II and III activities from sea urchin. On the basis of their sensitivity to α-amanitin, Pol I and II were counterparts of Pol A and B, whereas Pol III was later found[11] to be inhibited by α-amanitin at a much higher concentration, similarly to Pol C activities found by others[8].

The subunit composition of purified Pol A and B[8,12], and subsequently those of Pol I, II and III[13], confirmed that Pol I and II are identical to Pol A and B, and showed that Pol III is a different enzyme. Others also found Pol A, B and C in yeast, showing similar subunit structures (14, 12 and 17 subunits for Pol A, B and C, respectively). This structural complexity was in marked contrast with the single four-subunit bacterial enzyme, in keeping with the idea that the molecular mechanisms controlling transcription would prove to be much more complex in eukaryotes than in prokaryotes.

Using α-amanitin, studies with cultured cells and isolated nuclei showed that preribosomal RNA, cellular and viral pre-mRNA, and pre-tRNA and 5S RNA are synthesized by Pol I (A), II (B) and III (C), respectively[8,14]. However, none of our purified RNA polymerases, notably Pol B, could initiate transcription on intact double-stranded DNA[8]. Because chromatin, rather than naked DNA, was probably the template *in vivo*, we prepared chromatin as gently as possible. The unexpected outcome was our contribution to the elucidation of the structure of chromatin and of its repeating unit, which we named the 'nucleosome'[15,16]. However, chromatin was transcribed no better than intact DNA.

Because protein-coding genes represent a small fraction of vertebrate genomes, these failures could have been due to the scarcity of Pol B promoters. However, intact adenovirus-2 (Ad2) DNA, known to be tran-

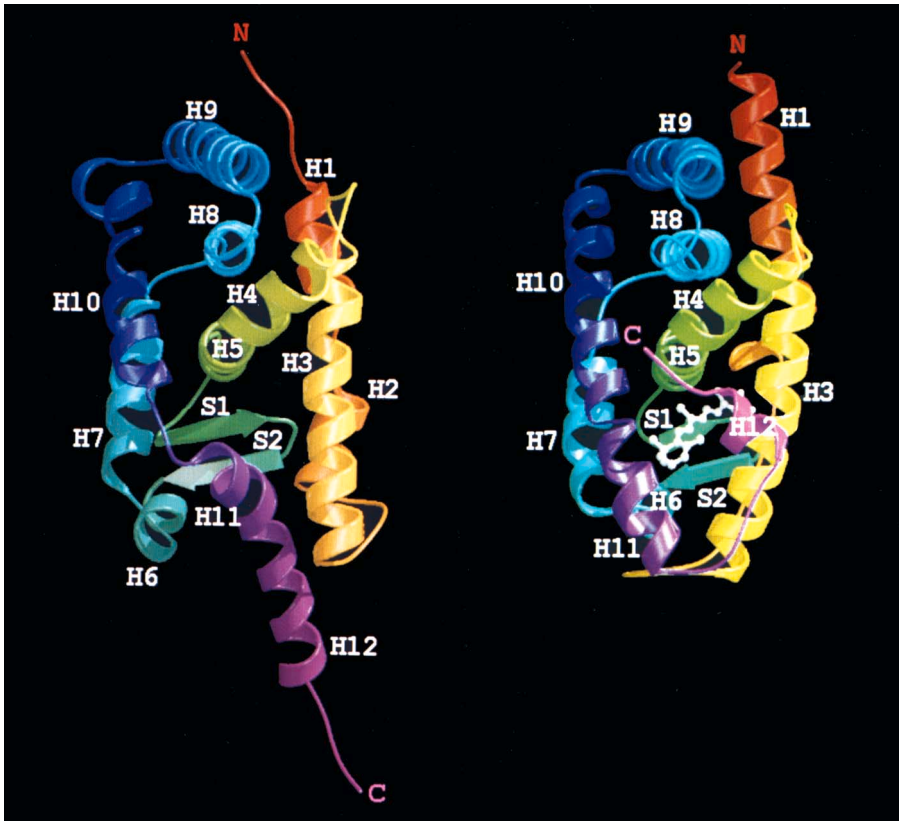

Figure 1 Ribbon diagrams of the unliganded human RXRα (left) and of the human RARγ bound to all-trans retinoic acid in a ball-and-stick representation (right). Note the large conformational shift elicited by the ligand[71–74]. H1 to H12, individual α-helices; S1 and S2, β-strands. Figure courtesy of J.P. Renaud and W. Bourguet.

scribed by Pol B *in vivo*, was no better as a template. Because similar results were reported by others[8], it was clear that either purified Pol B lacked bacterial σ-like factors (lost during purification) or some factors not bound to Pol B and with no counterparts in prokaryotes might be required in eukaryotes for transcription of promoters of protein-coding genes.

What was necessary to look for such factors were short DNA templates containing bona fide Pol B promoters. Because none were available in 1974, we were in the situation that I had feared several years earlier. What was desperately needed was a genetic toolbox that would allow us to obtain analyzable amounts of well-defined DNA fragments from any genome, in order to sequence DNA, to identify accurately RNA 5′ ends, to create site-directed mutations and so on. With the advent of the technological advances that fueled the genetic revolution, everything that we had thought would never be possible suddenly turned into reality. The excitement was terrific, because we could foresee the day when the mysteries underlying the control of genetic programs in higher eukaryotes would be unraveled at the molecular level.

Complexity of promoter regions of protein-encoding genes

Because we were interested in hormonally controlled transcription, we promptly cloned the estrogen- and progesterone-inducible ovalbumin and conalbumin genes to look for promoter elements around their 5′ ends[17,18]. Incidentally, while cloning the ovalbumin gene in 1977, we made the remarkable discovery that DNA sequences encoding the protein are split; this was the first evidence that the amino acid coding sequences of eukaryotic genes could be interrupted[19], whereas Phil Sharp's group and a Cold Spring Harbor Laboratory team had shown a few months earlier that an Ad2 gene was split in its 5′ leader sequence. Using deletion and single-point mutants of cloned DNA fragments of ovalbumin, conalbumin and Ad2 major late genes, we showed that the Goldberg-Hogness TATA homology was crucial for fixing initiation of transcription at the start sites, both *in vitro*[17,18,20–22] using a soluble system[23] dependent on purified Pol B, and *in vivo* using *Xenopus* oocytes[24]. Furthermore, we found that a partially purified initiation factor (BTF1; TFIID in Roeder's nomenclature) required *in vitro* stably bound the TATA element to form a preinitiation complex in the absence of Pol B[25]. This was markedly at variance with the prokaryotic promoter paradigm, whereby RNA polymerase holoenzymes (which include σ-initiation factors) recognize and bind on their own to proximal promoter elements. Further studies showed that the evolutionarily conserved TATA-binding protein (TBP)[26–28] was stably associated within the TFIID complex with several polypeptides (up to 14) that Tjian named TBP-associated factors (TAFs)[29]. Moreover, studies from our laboratory[30–32] and others indicated that distinct populations of TFIID complexes might exist, possibly mediating the effect of different transcriptional activators bound to further distal promoter elements (see later). Again, none of this had a precedent in prokaryotes.

One or several of a variety of upstream promoter elements were found at variably located positions between positions –40 and –110. Most were present in several genes, but some might be gene specific. They were found to be required for efficient transcription *in vitro* and *in vivo*, as initially shown by us for the Ad2 major late upstream element and its cognate factor UEF/USF[33–35], and by Tjian and co-workers and ourselves for the Sp1 factor that binds the GC-rich upstream elements of SV40 early and late promoters and a number of cellular genes[36–40]. Such a variety of upstream promoter elements scattered over 70 base pairs (bp) upstream from the start site had no precedent in bacteria.

The totally unexpected finding, however, came with the functional dissection of the SV40 early promoter achieved in 1979–1981 by Christophe Benoist, who discovered that the selective deletion of a remote sequence that included the 72-bp tandem repeat located (at position –116 to –261) upstream from the GC-rich upstream elements markedly decreased transcription initiated at the SV40 early gene start sites, even though the TATA element and GC-rich elements were intact[41,42]. In contrast to the TATA box, which was required *in vitro*[43] and *in vivo*[42] to fix initiation at SV40 early start sites, the 72-bp repeat element had no apparent effect on transcription *in vitro*. We quickly discovered that the SV40 72-bp repeat had several amazing properties, including the capacity to enhance transcription from heterologous promoters and to act in either orientation and over a broad range of distances upstream and downstream of the stimulated pro-

moter[44,45]. While using a SV40 vector to study mechanisms controlling the expression of a rabbit β-globin gene, Schaffner and coworkers independently found similar effects of the 72-bp repeat, which they named an 'enhancer'[46]. Enhancer elements with similar properties were found soon after in cellular genes, notably in the immunoglobin heavy chain gene locus where it showed tissue specificity[47,48].

Because comparing various enhancers did not reveal any obvious sequence homologies and their mode of action was still a matter of intense but sterile speculation two years after their discovery, we made systematic deletions and point mutations throughout the enhancer region and investigated their effects[49]. The minimal SV40 enhancer included one 72-bp sequence and some 30 bp further upstream, and enhancement in HeLa cells could be ascribed to seven specific motifs that could overlap, whereas an additional motif was specifically involved in the enhancer activity in lymphoid cells[49,50]. Each of these motifs showed little activity on its own, whereas their association resulted in a 400-fold synergistic transcriptional enhancement. *In vitro* 'footprinting' in HeLa and lymphoid cells showed that each motif could bind a protein. For each motif, this binding was prevented by mutations detrimental to enhancer activity *in vivo*, indicating that the bound proteins could correspond to *trans*-acting factors mediating enhancer activity[50–52]. Because the different SV40 enhancer motifs were found in various assortments in other viral and cellular enhancers[49], all of these observations led us to suggest[49,52] that the activity of an enhancer in a given cell type could result from both the nature of its constituent motifs and the presence or absence of its cognate *trans*-acting factors in that cell type. The general principles we had learned from the modular organization of the prototypic SV40 enhancer clearly showed the combinatorial possibilities of controlling transcription with a limited number of factors bound to enhancers in a cell type–specific manner.

Thus, several of the actors most likely involved in combinatorial control of transcription initiation in higher eukaryotes had been identified 20 years ago with the discovery of (i) the multisubunit RNA Pol B (II in Roeder's nomenclature) dedicated to transcription of protein-encoding genes (ii) several TAFs present in preinitiation TFIID complexes bound to the TATA element (iii) multiple gene–restricted upstream promoter elements and their possibly cell-specific cognate activators and (iv) modular enhancers and their motif-specific cell type–restricted bound factors. What remained unknown was the molecular mechanism by which these factors, notably those bound to enhancers, could trigger initiation of transcription. While we were looking in 1983 for enhancers simpler than that of SV40 to tackle this problem, it became apparent that a glucocorticoid response element present in the mouse mammary tumor virus (MMTV) long terminal repeat[53,54] had the characteristic properties of an enhancer element to which the glucocorticoid receptor (GR) could bind to trigger transcription from a heterologous promoter[54,55]. In other words, the GR might be a ligand-dependent enhancer-binding regulatory protein that, in contrast to the factors bound to the different SV40 enhancer motifs, could be active on its own. Assuming that all steroid hormone receptors might be similar to the GR, we looked at them as almost ideal objects from which we could not only reveal how enhancer-binding activators trigger transcription, but also how nuclear receptors specifically recognize their cognate *cis*-acting DNA response elements and become transcriptionally active upon binding of their ligands. Because such studies required reverse genetic approaches, the first step was the cloning of receptor cDNAs. In view of our long-standing interest in estrogen action, and because it was clear that GR cloning was underway in other laboratories, we set out to clone the ER cDNA.

Lessons from our functional dissection of the estrogen receptor

Using antibodies prepared by Geoff Greene and Elwood Jensen against purified human ER (known today as ERα), as well as synthetic oligonucleotide probes derived from its amino acid sequence, we[56] isolated and sequenced ER cDNA clones in 1985, one of which contained the entire ER open reading frame[57,58]. Shortly afterward we cloned the chicken ER cDNA[59]. A comparison of these two cDNAs revealed six regions (A–F from the N-terminal to the C-terminal ends) of variable evolutionary conservation. Two of them (C and E) were highly conserved and similarly present in human GR, which had just been cloned in Ron Evans' lab and, more surprisingly, in the v-erb-a product of the avian erythroblastosis virus, which provided the first indication that the ER and GR are members of a larger gene family.

We showed *in vitro* that region E was the ligand-binding domain (LBD) and region C was the DNA-binding domain (DBD)[60,61], and we unequivocally confirmed this *in vivo* by constructing a chimeric receptor in which we swapped region C of ER for that of GR. Notably, the chimera could bind estradiol but could not activate a chimeric estradiol-responsive reporter gene (*vit-tk-CAT*), whereas it activated a glucocorticoid-responsive gene in the presence of estradiol, but not in the presence of a glucocorticoid. In a reciprocal construct, the chimeric receptor could, in the presence of a glucocorticoid, activate the estrogen-responsive reporter *vit-tk-CAT*, but not the chimeric glucocorticoid-responsive reporter gene *MMTV-CAT*[62,63]. These data not only showed that steroid hormone receptors (and presumably other members of the family) have a modular structure in which the DBD and the LBD can function independently, but importantly, also indicated how such chimeric receptors can be used to identify ligands for any novel member of the nuclear receptor (NR) family: hooking the putative LBD of a new receptor to the DBD of either the ER or the GR should make its ligand activate an estrogen- or a glucocorticoid-responsive reporter gene, respectively. The value of this approach was quickly illustrated by the independent discovery of the first retinoic acid receptor (RARα) by us[64] and Ron Evans' lab[65].

By 1987 the stage had been set for bringing the NR superfamily to completion[66]. Furthermore, by 1990 we had shown that the ER has two distinct synergistic activation functions (the ligand-independent AF-1 located in the A/B region and the ligand-dependent AF-2 encompassing the LBD) and had proposed that their transcriptional activity is mediated by intermediary factors (TIFs)[67,68]. This proposal was confirmed by the discovery in 1995–1996 of the first two bona fide p160 coactivators by O'Malley and colleagues (SRC-1)[69] and our group (TIF-2)[70]. Because these coactivators were found to directly or indirectly (through CBP/p300) acetylate the nucleosome core histones, these results established that an activation function, such as ER AF-2, could stimulate transcription through the recruitment of factors remodeling the chromatin template and through interaction with or recruitment of factors belonging to the transcription apparatus such as TAFs[30–32], thereby enormously increasing the diversity of the transcriptional response to a hormonal signal. Moreover, our crystallographic studies of bound and unbound LBDs of retinoid X receptor-α (RXRα) and RARγ[71–74] (**Fig. 1**), performed in collaboration with Dino Moras' group in our Institute, revealed that the mechanism involved in the activation of AF-2 upon ligand binding corresponded to a conformational shift of the LBD, creating an

interacting surface to which factors such as the coactivators could bind, whereas corepressors would be released[73,75].

Finally, in collaboration with Shigeaki Kato and co-workers[76], we showed that the ER AF-1 activity is enhanced by phosphorylation of a serine residue, thereby establishing that signals transduced by membrane-associated receptors through kinase cascades can modulate the activity of NRs[77]. Thus, nuclear receptors are highly sophisticated molecular switches that are capable of responding to the two main cellular signaling pathways by simultaneously transducing their cognate ligand signal directly as well as signals received by membrane-associated receptors.

RAR and RXR families: an extraordinary combinatorial complexity

It is beyond the scope of this commentary to summarize all of the work that we have done since the discovery of the three RAR isotype genes (α, β and γ, encoding altogether eight main isoforms) and the three RXR isotype genes (α, β and γ, encoding altogether six main isoforms). The demonstration by us and others[75,77,78] that they can transduce the retinoid signal as RAR-RXR heterodimeric functional units, and furthermore that RXRs also heterodimerize with TRs, VDR, PPARs, LXRs, FXR, PXR, Nurr1 and few other NRs, strikingly illustrated the concept of combinatorial control of gene expression in higher eukaryotes. Our *in vitro* reconstruction, from purified components, of a ligand-dependent transcription system that mimics transactivation by retinoids *in vivo* and is dependent on remodeling of the chromatin template[79,80], paves the way to the full elucidation, at the molecular level, of the problem we raised 40 years ago: how could the binding of a hormone to its cognate nuclear receptor act positively on initiation of transcription? Remembering where I was 40 years ago, I am every day astonished that we succeeded in demonstrating the reality of this control during mouse development[81,82] and in homeostasis[83] using conventional germline site-directed knockouts[84] and, more recently, spatiotemporally controlled site-directed somatic mutations[83]. I am even more amazed that all of this happened by tinkering, to borrow the beautiful image from François Jacob. Is it, however, so extraordinary? After all, the transcription machinery required for activation of gene expression by nuclear receptors was already present in yeast more than one billion years ago[85], even though there are no nuclear receptors, no cognate ligands and no p160 coactivators in this unicellular eukaryote.

ACKNOWLEDGMENTS

I thank my students, my postdocs and all of my collaborators for their enthusiastic contributions over all these years, G. Richards for a critical reading of the manuscript, and the French and foreign institutions that have supported my laboratory. Last, but not least, I am grateful to all members of my own family who have always been "super" over all these years.

1. Weiss, S.B. & Gladstone, L. A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. *J. Am. Chem. Soc.* **81**, 4118–4119 (1959).
2. Ramuz, M., Doly, J., Mandel, P. & Chambon, P. A soluble DNA-dependent RNA polymerase in nuclei of non-dividing animal cells. *Biochem. Biophys. Res. Commun.* **19**, 114–120 (1965).
3. Chambon, P., Weill, J.D., Doly, J., Strosser, M.T. & Mandel, P. On the formation of a novel adenylic compound by enzymatic extracts of liver nuclei. *Biochem. Biophys. Res. Commun.* **25**, 638–643 (1966).
4. Ménissier-de Murcia, J. *et al.* Functional interaction between PARP-1 and PARP-2 in chromosome stability and embryonic development in mouse. *EMBO J.* **22**, 2255–2263 (2003).
5. Amé, J-C., Spenlehauer, C. & de Murcia, G. The PARP superfamily. *Bioessays* **26**, 882–893 (2004).
6. Weill, J.D., Busch, S., Chambon, P. & Mandel, P. The effect of estradiol injections upon chicken liver nuclei RNA polymerase. *Biochem. Biophys. Res. Commun.* **10**, 122–126 (1963).
7. Stirpe, F. & Fiume, L. Studies on the pathogenesis of liver necrosis by α-amanitin. *Biochem. J.* **105**, 779–782 (1967).
8. Chambon, P. Eukaryotic nuclear RNA polymerases. *Annu. Rev. Biochem.* **44**, 613–638 (1975).
9. Roeder, R.G. & Rutter, W.J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**, 234–237 (1969).
10. Kédinger, C., Gniazdowski, M., Mandel, J.L., Gissinger, F. & Chambon, P. α-Amanitin: a specific inhibitor of one of two DNA dependent RNA polymerase activities from calf thymus. *Biochem. Biophys. Res. Comm.* **38**, 165–171 (1970).
11. Weinmann, R. & Roeder, R.G. Role of DNA-dependent RNA polymerase III in transcription of the tRNA and 5S RNA genes. *Proc. Natl Acad. Sci. USA* **71**, 1790–1794 (1974).
12. Chambon, P. *et al.* Purification and properties of Calf thymus DNA-dependent RNA polymerases A and B. *Cold Spring Harb. Symp. Quant. Biol.* **35**, 693–707 (1970).
13. Sklar, V.E.F., Schwartz, L.B. & Roeder, R.G. Distinct molecular structure of nuclear class I, II and III DNA-dependent RNA polymerases. *Proc. Natl Acad. Sci. USA* **72**, 348–352 (1975).
14. Roeder, R.G. Eukaryotic nuclear RNA polymerases. in *RNA Polymerase* (R. Losick & M. Chamberlin, eds.) 285–329 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1979).
15. Oudet, P., Gross-Bellard, M. & Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**, 281–300 (1975).
16. Oudet, P. & Chambon, P. Seeing is believing. *Cell* **S116**, S79–S80 (2004).
17. Gannon, F. *et al.* Organisation and sequences at the 5′ end of a cloned complete ovalbumin gene. *Nature* **278**, 423–434 (1979).
18. Cochet, M. *et al.* Organisation and sequence studies of the 17-piece conalbumin gene. *Nature* **282**, 567–574 (1979).
19. Breathnach, R., Mandel, J.L. & Chambon, P. The ovalbumin gene is split in chicken DNA. *Nature* **270**, 314–319 (1977).
20. Wasylyk, B., Kédinger, C., Corden, J., Brison, O. & Chambon, P. Specific in vitro initiation of transcription on conalbumin and ovalbumin genes and comparison with adenovirus-2 early and late genes. *Nature* **285**, 388–390 (1980).
21. Corden, J. *et al.* Promoter sequences of eukaryotic protein-coding genes. *Science* **209**, 1406–1414 (1980).
22. Wasylyk, B. *et al.* Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in TATA box homology sequence. *Proc. Natl Acad. Sci. USA* **77**, 7024–7028 (1980).
23. Wu, G-J. Adenovirus DNA-directed transcription of 5.5S RNA in vitro. *Proc. Natl Acad. Sci. USA* **75**, 2175–2219 (1978).
24. Grosschedl, R., Wasylyk, B., Chambon, P. & Birnstiel, M.L. Point mutation in the TATA box curtails expression of sea urchin H2A histone gene in vivo. *Nature* **294**, 179–181 (1981).
25. Davison, B.L., Egly, J-M., Mulvihill, E.R. & Chambon, P. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature* **301**, 680–686 (1983).
26. Cavallini, B. *et al.* Cloning of the gene encoding the yeast protein BTF1Y that can substitute for the human TATA box factor. *Proc. Natl Acad. Sci. USA* **86**, 9803–9807 (1989).
27. Buratowski, S., Hahn, S., Sharp, P.A. & Guarente, L. Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334**, 37–42 (1988).
28. Cavallini, B. *et al.* A yeast activity can substitute for the HeLa cell TATA box factor in vitro. *Nature* **334**, 77–80 (1988).
29. Dynlacht, B.D., Hoey, H. & Tjian, R. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**, 563–576 (1991).
30. Brou, C. *et al.* Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J.* **12**, 489–499 (1993).
31. Jacq, X. *et al.* Human $TAF_{II}30$ is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* **79**, 107–117 (1994).
32. Mengus, G., May, M., Carré, L., Chambon, P. & Davidson, I. Human TAF_{II}^{135} potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Genes Dev.* **11**, 1381–1395 (1997).
33. Hen, R., Sassone-Corsi, P., Corden, J., Gaub, M.P. & Chambon, P. Sequences upstream from the TATA box are required *in vivo* and *in vitro* for efficient transcription from the Adenovirus-2 major late promoter. *Cell* **43**, 165–175 (1985).
34. Sawadogo, M. & Roeder, R.G. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**, 165–175 (1985).
35. Miyamoto, N.G., Moncollin, V., Egly, J.M. & Chambon, P. Specific interaction between a transcription factor and the upstream element of the Adenovirus-2 major late promoter. *EMBO J.* **4**, 3563–3570 (1985).
36. Dynan, W.S. & Tjian, R. The promoter specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**, 79–87 (1983).
37. Everett, R.D., Baty, D. & Chambon, P. The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res.* **11**, 2447–2464 (1983).
38. Barrera-Saldana, H. *et al.* All six GC-motifs of the SV40 early upstream element contribute to promoter activity *in vivo* and *in vitro*. *EMBO J.* **4**, 3839–3849 (1985).
39. Gidoni, D. *et al.* Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. *Science* **230**, 511–517 (1985).
40. Kadonaga, J.T., Carner, K.R., Masiarz, F.R. & Tjian, R. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**, 1079–1090 (1987).
41. Benoist, C. & Chambon, P. Deletions covering the putative promoter region of early mRNAs of simian virus 40 do not abolish T-antigen expression. *Proc. Natl Acad. Sci. USA* **77**, 3865–3869 (1980).
42. Benoist, C. & Chambon, P. *In vivo* sequence requirements of the SV40 early promoter region. *Nature* **290**, 304–310 (1981).
43. Mathis, D. & Chambon, P. The SV40 early region TATA box, but not the upstream sequences, is required for accurate in vitro initiation of transcrip-

tion. *Nature* **290**, 310–316 (1981).
44. Moreau, P. *et al.* The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* **9**, 6047–6068 (1981).
45. Wasylyk, B., Wasylyk, C., Augereau, P. & Chambon, P. The SV40 72bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. *Cell* **32**, 503–514 (1983).
46. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
47. Banerji, J., Olson, L. & Schaffner, W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglublin heavy chain genes. *Cell* **33**, 729–740 (1983).
48. Gillies, S.D., Morrison, S.L., Oi, V.T. & Tonegawa, S. A tissue specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**, 717–728 (1983).
49. Zenke, M. *et al.* Multiple sequence motifs are involved in SV40 enhancer function. *EMBO J.* **5**, 387–397 (1986).
50. Nomiyama, H., Fromental, C., Xiao, J.H. & Chambon, P. Cell-specific activity of the constituent elements of the SV40 enhancer. *Proc. Natl Acad. Sci. USA* **84**, 7881–7885 (1987).
51. Wildeman, A. *et al.* Specific protein binding to the SV40 enhancer *in vitro*. *Mol. Cell. Biol.* **6**, 2098–2105 (1986).
52. Davidson, I. *et al.* Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. *Nature* **323**, 544–548 (1986).
53. Scheidereit, C., Geisse, S., Westphal, H.M. & Beato, M. The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumour virus. *Nature* **304**, 749–752 (1983).
54. Yamamoto, K. Steroid Receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* **19**, 209–252 (1985).
55. Chandler, V.L., Maler, B.A. & Yamamoto, K.R. DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell* **33**, 489–499 (1983).
56. Walter, P. *et al.* Cloning of the human estrogen receptor cDNA. *Proc. Natl Acad. Sci. USA* **82**, 7889–7893 (1985).
57. Green, S. *et al.* The human oestrogen cDNA: sequence, expression and homology to v-erbA. *Nature* **320**, 134–139 (1986).
58. Greene, G.L. *et al.* Sequence and expression of human estrogen receptor complementary DNA. *Science* **231**, 1150–1154 (1986).
59. Krust, A. *et al.* The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. *EMBO J.* **5**, 891–897 (1986).
60. Kumar, V., Green, S., Staub, A. & Chambon, P. Localisation of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor. *EMBO J.* **5**, 2231–2236 (1986).
61. Kumar, V. *et al.* Functional domains of the human estrogen receptor. *Cell* **51**, 941–951 (1987).
62. Green, S. & Chambon, P. Oestradiol induction of a glucocorticoid-responsive gene by a chimeric receptor. *Nature* **325**, 75–78 (1987).
63. Kumar, V. *et al.* Functional domains of the human estrogen receptor. *Cell* **51**, 941–951 (1987).
64. Petkovich, M., Brand, N.J., Krust, A. & Chambon, P. A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* **330**, 444–450 (1987).
65. Giguère, V., Ong, E.S., Segui, P. & Evans, R.M. Identification of a receptor for the morphogen retinoic acid. *Nature* **330**, 624–629 (1987).
66. Laudet, V. & Gronemeyer, H. *The Nuclear Receptor Facts Book* (Academic, London, 2002).
67. Tora, L. *et al.* The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell* **59**, 477–487 (1989).
68. Tasset, D., Tora, L., Fromental, C., Scheer, E. & Chambon, P. Distinct classes of transcriptional activating domains function by different mechanisms. *Cell* **62**, 1177–1187 (1990).
69. Onate, S.A., Tsai, S.Y., Tsai, M.J. & O'Malley, B. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* **270**, 1354–1357 (1995).
70. Voegel, J.J. Heine, M.J.S. Zechel, C. Chambon, P. & Gronemeyer, H. TIF2, a 160-kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. *EMBO J.* **15**, 3667–3675 (1996).
71. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. Crystal structure of the ligand binding domain of the human nuclear receptor RXRα. *Nature* **375**, 377–382 (1995).
72. Renaud, J-P. *et al.* Crystal structure of the RAR-γ ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681–689 (1995).
73. Wurtz, J-M. *et al.* A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.* **3**, 87–94 (1996).
74. Egea, P.F. *et al.* Crystal structure of the human RXRα ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *EMBO J.* **19**, 2592–2601 (2000).
75. Chambon, P. A decade of molecular biology of retinoic acid receptors. *FASEB J.* **10**, 940–954 (1996).
76. Kato, S. *et al.* Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science* **270**, 1491–1494 (1995).
77. Rochette-Egly, C. & Chambon, P. F9 Embryocarcinoma cells: a cell autonomous model to study the functional selectivity of RARs and RXRs in retinoid signaling. *Histol. Histopathol.* **16**, 909–922 (2001).
78. Leid, M., Kastner, P. & Chambon, P. Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends Biochem. Sci.* **17**, 427–433 (1992).
79. Dilworth, F.J., Fromental-Ramain, C., Yamamoto, K. & Chambon, P. ATP-driven chromatin remodeling activity and histone acetyltransferases act sequentially during transactivation by RAR/RXR *in vitro*. *Mol. Cell* **6**, 1049–1058 (2000).
80. Dilworth, F.J. & Chambon, P. Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene Rev.* **20**, 3047–3054 (2001).
81. Kastner, P., Mark, M. & Chambon, P. Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* **83**, 859–869 (1995).
82. Mark, M. & Chambon, P. Functions of RARs and RXRs in vivo: Genetic dissection of the retinoid signaling pathway. *Pure Appl. Chem.* **75**, 1709–1732 (2003).
83. Chapellier, B. *et al.* Physiological and retinoid-induced proliferations of epidermal basal keratinocytes are differently controlled. *EMBO J.* **21**, 3402–3413 (2002).
84. Thomas, K.R. & Capecchi, M.R. Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in the midbrain and cerebellar development. *Nature* **346**, 847–850 (1990).
85. Metzger, D., White, J.H. & Chambon, P. The human oestrogen receptor functions in yeast. *Nature* **334**, 31–36 (1988).

Pierre Chambon is at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (CNRS, INSERM, ULP, Collège de France) and the Institut Clinique de la Souris, B.P.10142, 67404, Illkirch, Strasbourg, France.
e-mail: chambon@igbmc.u-strasbg.fr