The EMBO Journal vol.11 no.2 pp.629-642, 1992

Structure, localization and transcriptional properties of two classes of retinoic acid receptor α fusion proteins in acute promyelocytic leukemia (APL): structural similarities with a new family of oncoproteins

Philippe Kastner[1], Aymée Perez[1], Yves Lutz, Cécile Rochette-Egly, Marie-Pierre Gaub, Béatrice Durand, Michel Lanotte[2], Roland Berger[2] and Pierre Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes, CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique, INSERM, Institut de Chimie Biologique, Faculté de Médecine 11, rue Humann, 67085 Strasbourg Cédex, and [2]Unité 301 de l'INSERM, SDI No 15954.I CNRS, Centre Hayem and IGM, and Laboratoire Central d'Hematologie, Hôpital Saint-Louis, Paris, France

[1]Should be considered as equal first authors

Communicated by P.Chambon

Acute promyelocytic leukemia (APL) is due to a chromosomal t(15;17) translocation which involves a novel human gene, Myl, (also named PML) and the retinoic acid (RA) receptor α (RAR-α) gene. We report here the characterization of Myl and of the reciprocal MylRAR (PMLRAR) and RARMyl (RARPML) fusion transcripts which are found in two classes of APL patients. Myl displays similarities with a new family of proteins of which some members are fused to proto-oncogenes in the transforming proteins RFP-ret and T18. The speckled nuclear localization of Myl, as well as its sequence homology with the 52 kDa component of the RO/SSA ribonucleoprotein particle, suggest that Myl may be present in a ribonucleoprotein complex. In contrast to both Myl and RAR-α whose localization is essentially nuclear in the presence or absence of RA, MylRAR which is largely cytoplasmic in the absence of RA appears to be translocated to the nucleus in the presence of RA. Myl and MylRAR can associate *in vitro* and this association is mediated by a coiled coil in the Myl sequence. *In vivo* this association results in a colocalization of Myl and MylRAR which is identical to that of MylRAR alone. Studies of activation of transcription from the promoters of several RA target genes indicate that MylRARs have altered transcription activation properties when compared with RAR-α. Most notably, MylRAR represssses markedly the activity of some RA target promoters in the absence of RA. Western blot analyses of patient samples show that MylRAR is expressed to a much higher level than wild type RAR-α originating from the normal allele. Taken together, these results suggest that MylRAR may interfere in a dominant manner with both Myl and RAR functions.

Key words: APL/coiled coil/Myl (PML)/RAR-α/MylRAR (PMLRAR) fusion/RARMyl (RARPML) fusion

Introduction

Chromosomal abnormalities are frequently associated with malignant diseases. In a number of instances, specific chromosomal translocations have been characterized, which generate fusion genes encoding proteins with oncogenic properties (Cleary, 1991; Sawyers *et al.*, 1991). A specific t(15;17) reciprocal translocation is the hallmark of human acute promyelocytic leukemia (APL) (Rowley *et al.*, 1977). The localization of the human retinoic acid receptor α (RAR-α) gene in the region of chromosome 17 which is involved in this translocation, led us to speculate that disruption of this gene may be related to APL etiology (Mattei *et al.*, 1988). It has been shown subsequently that the chromosome 17 breakpoint lies in fact within the RAR-α gene (Borrow *et al.*, 1990; de Thé *et al.*, 1990; Alcalay *et al.*, 1991; Chen *et al.*, 1991) and that, in APL cells, transcripts are produced from a fusion between Myl, a novel gene on chromosome 15, and the RAR-α gene (de Thé *et al.*, 1990; Longo *et al.*, 1990; Warrel *et al.*, 1991). This finding is especially interesting, since APL patients undergo complete remission after a few weeks of treatment with retinoic acid (Huang *et al.*, 1988; Castaigne *et al.*, 1990; Chomienne *et al.*, 1989; Warrel *et al.*, 1991; Clarkson, 1991, for review).

Retinoic acid (RA) is a vitamin A derivative which exerts profound effects on vertebrate development and on cell growth and differentiation (Brockes, 1989 for a review). Three RAR genes, RAR-α, β and γ have been characterized and it has been shown that several RAR isoforms can be generated from each gene subtype (either α, β or γ) by differential use of two promoters and alternative splicing (Giguere *et al.*, 1987; Petkovich *et al.*, 1987; Brand *et al.*, 1988; Zelent *et al.*, 1989, 1991; Kastner *et al.*, 1990; Leroy *et al.*, 1991). RARs belong to the superfamily of nuclear steroid/thyroid hormone receptors which act as ligand-inducible transcription factors modulating the expression of target genes (Evans, 1988; Green and Chambon, 1988 for reviews). The amino acid sequence of RARs has been divided into six regions (A-F) based on different degrees of conservation, with regions C and E being highly conserved (Kastner *et al.*, 1991, for review). Region C contains the zinc finger DNA binding domain, whereas region E corresponds to a multifunctional region containing both the RA binding domain, a RA-inducible transcription activation function and possibly an interface involved in protein-protein interaction (Glass *et al.*, 1990). Regions B, D and F are less or not conserved when comparing the three RAR subtypes within a given species, whereas they are highly conserved for a given subtype across species. Furthermore, the N-terminal A region which is isoform specific is also conserved across species for a given isoform of either RAR-α, β or γ. These region A differences may correspond to cell- and promoter-specific transcriptional activation properties of the various RAR isoforms (S.Nagpal and P.Chambon, unpublished results). The possible functions of B, D and F regions is unknown.

We report here the cloning of Myl and the characterization of two sets of MylRAR and RARMyl transcripts and proteins that occur in two classes of APL patients. Structural features of the Myl protein, the intracellular localization of Myl,

MylRAR and RARMyl, the transactivation properties of the MylRAR chimeras on RA target genes, and the association between Myl and MylRAR, provide the basis for speculating on how these fusion proteins may be involved in APL etiology.

Results

Cloning of Myl (PML) cDNA

The point of fusion in the Myl−RAR-α transcript which was found in the APL-derived cell line NB4 (Lanotte *et al.*, 1991), is located at the splicing junction between the exons encoding the A and B regions of RAR-α (de Thé *et al.*, 1990; see RAR-α in Figure 2). We first used anchored PCR to clone 1798 nucleotides of Myl 5′ cDNA sequences corresponding to RNA sequences located upstream of the RAR-α B region in Myl−RAR RNA present in the NB4 cell line (Figure 1, sequence located upstream of the open triangle designated A). The 3′ part of the Myl cDNA was cloned by PCR using Myl-specific primers and oligo(dT)-containing primers. The total length of the cloned Myl cDNA is 2199 bp (Figure 1 and data not shown), which may correspond to the smallest Myl transcript species detected on Northern blots (de Thé *et al.*, 1990).

An open reading frame (ORF) conceptually encoding a 633 amino acid protein (calculated molecular weight 70 028 Daltons) is present in this Myl sequence (nucleotides 142−2040, Figure 1). Although the reading frame remains open up to the 5′ end of the cDNA, we could not isolate any further 5′-extended cDNA by anchored PCR, and the present sequence may therefore correspond to a nearly full length Myl cDNA. Furthermore, the similar size of cloned and endogenous MylRAR proteins (see below) supports the idea that the present cDNA sequence is not missing the 5′ end of the Myl coding sequence, and hence, that the first AUG (nucleotides 142−144) could be the Myl initiation codon. Both the first and the second (Met23) AUG are located in a sequence context acceptable for initiation codons (Cavener and Ray, 1991). Since *in vitro* transcribed and translated Myl mRNA yields a doublet protein band (data not shown), both AUGs may in fact be used as initiation codons.

We also identified two shorter cDNA isoforms, presumably lacking specific Myl exons (nucleotides 219−278 and 1396−1539, respectively; see Figure 1, sequences in brackets, and data not shown). These putative alternative splicing events would generate Myl proteins deleted for amino acids 27−46 and 419−466, named hereafter Myl1 (PML1) and Myl2 (PML2) isoforms (see Myl in Figure 2).

Hybridization of Myl cDNA to a Southern blot containing DNA from a panel of human−hamster cell hybrids (Bios Blot, Bios Corporation; data not shown) confirmed that the Myl gene is located on chromosome 15 (de Thé *et al.*, 1990).

Different MylRAR (PMLRAR) transcripts in two classes of APLs

The above results indicate that the cloned Myl−RAR-α fusion protein present in NB4 cells, named hereafter MylRAR-A (PMLRAR-A), should possess 552 Myl amino acids N-terminal to the RAR-α B region. Sequencing of NB4 cell cDNA synthesized by PCR using a Myl primer (nucleotides 1742−1762) and a RAR-α 3′-UTR primer (nucleotides 1597−2015, numbering as in Petkovich *et al.*, 1987), confirmed that the putative MylRAR-A contains the integrity of the RAR-α B−F regions and thus corresponds to a 955 amino acid long protein (mol. wt 105 815 Daltons, Figures 1 and 2, and data not shown). The presence in NB4 cells of RNA transcripts ~1 kb longer than RAR-α transcripts (de Thé *et al.*, 1990) is consistent with this conclusion. This altered pattern of RAR-α transcripts is similar to that found in a number of APL patients (Longo *et al.*, 1990; Warrel *et al.*, 1991), and we detected by PCR the MylRAR-A transcript in another APL patient, suggesting that the MylRAR-A form of MylRAR fusion transcript is not restricted to the NB4 cell line. In addition, sequencing of NB4 cell-derived cDNA (Figure 2b, primers M1 and R2), demonstrated the existence of MylRAR transcripts which were deleted for the same nucleotide sequence missing in the Myl2 isoform (see Figure 2b; lane 1). This alternative form of MylRAR-A protein will be referred to as Myl2RAR-A (PML2RAR-A) (see Figure 2). Myl1RAR-A (PML1RAR-A), the putative MylRAR-A isoform deleted for the Myl sequencing missing in the Myl1 isoform (Figure 2), has not yet been identified.

PCR-assisted analysis of RNA prepared from cells of seven other APL patients revealed MylRAR fusion products smaller than those found in NB4 cells (see Figure 2b, primers M1 and R2, lanes 3 and 4, and data not shown). Sequencing of these PCR-amplified cDNAs showed that the point of fusion with the RAR-α B region occurred at position 1324 in the Myl cDNA sequence (amino acid 394, see Figure 1, open triangle B). The deduced MylRAR fusion protein (named hereafter MylRAR-B or PMLRAR1B) is 797 amino acids long (mol. wt 89 288 Daltons), and comprises the first 394 amino acids of Myl and RAR-α B−F regions. Therefore, these patients represent a second class of APLs (class B), with MylRAR-B transcript and protein smaller than those found in the NB4-like class A patients, and likely to correspond to the shorter class of size-altered RAR-α transcripts reported previously in some APL patients (Longo *et al.*, 1990; Warrel *et al.*, 1991). The putative MylRAR-B cDNA isoform corresponding to the Myl1 isoform (Myl1RAR-B, see Figure 2a) has not yet been detected.

Reciprocal RARMyl (RARPML) transcripts in APL cells

We looked for the possible existence of fused RARα−Myl transcripts originating from the reciprocal t(15;17) translocation. PCR amplification with primers from the RAR-α1 5′-UTR and the Myl 3′ region (see Figure 2b, primers R1 and M2, respectively) yielded amplified cDNAs with RNA prepared from NB4 cells and class B APL patients, but not with RNA derived from promyelocytic leukemia HL60 cells (Breitman *et al.*, 1980) (Figure 2c and data not shown). Sequencing demonstrated that these cDNAs correspond to reciprocal RARMyl transcripts (Figure 2a). In NB4 cells (class A patients), the transcript conceptually encodes a 140 amino acid long protein (RARMyl-A or RARPML-A), mol. wt 14 958 Daltons) which fuses the A1 region of RAR-α1 to the 80 C-terminal amino acids of Myl, whereas the corresponding class B patient transcript encodes a 298 amino acid long protein (RARMyl-B or RARPML-B, mol. wt 31 484 Daltons) (Figure 2a, and data not shown). In class B patients, we found also RARMyl2-B transcripts in which Myl nucleotides 1396−1539 are deleted (Figures 1, 2a and b), conceptually encoding a 250 amino acid long protein (mol. wt 26 498 Daltons).

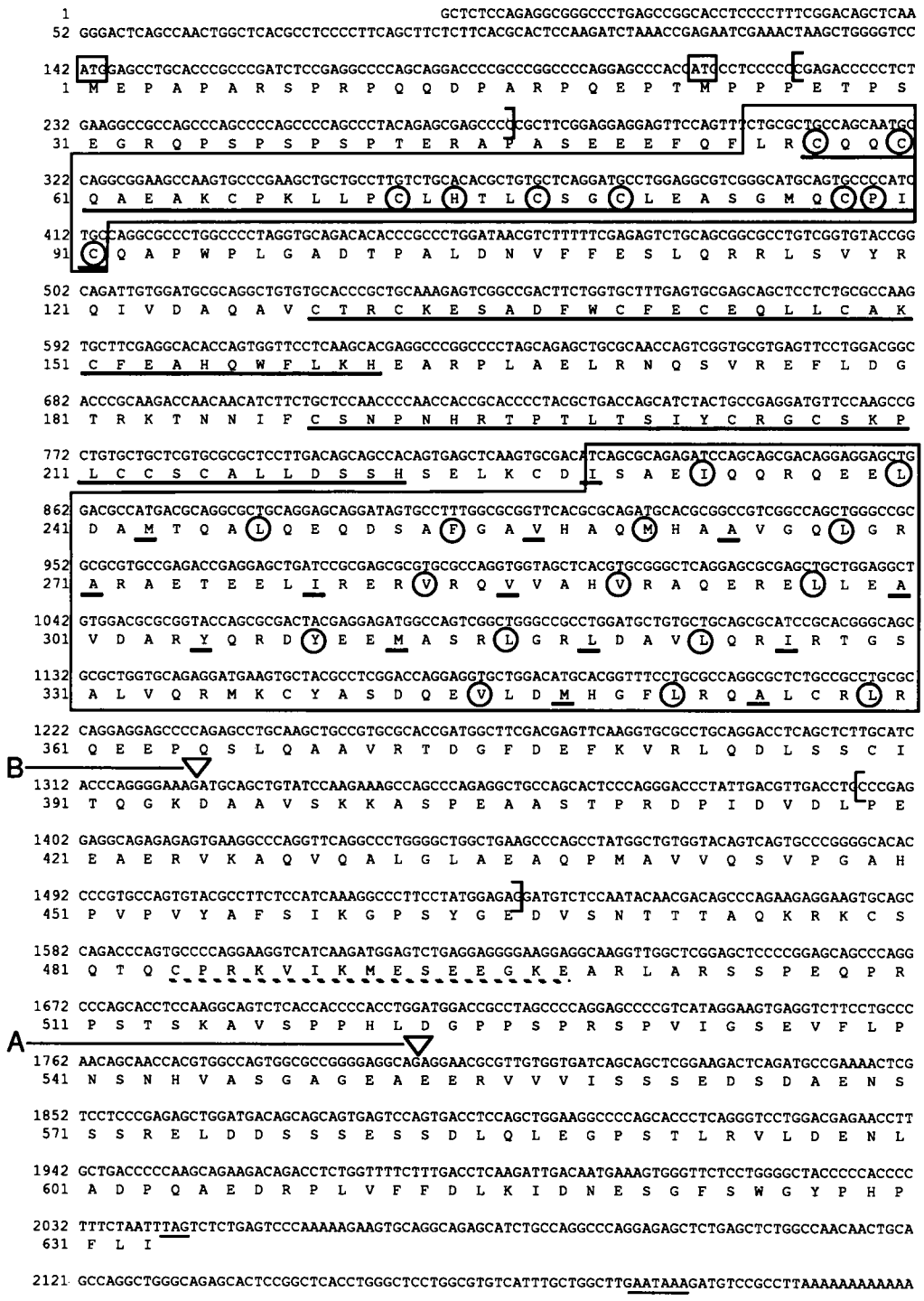

Fig. 1. Myl cDNA and amino acid sequences. The deduced amino acids are shown below their respective codons. The two regions (nucleotides 219–278 and 1396–1539, respectively) which are excluded by alternative splicing in Myl1 and Myl2 (see Figure 2a) are bracketed. The TAG stop codon and the polyadenylation signal are underlined. The open triangles A and B indicate the point where the Myl sequence is fused to RAR in class A and B APLs, respectively. The peptide used to generate Myl antibodies is underlined by a dashed line. The first boxed region corresponds to the first cysteine-rich motif (Figure 3a); within it, the conserved residues are circled. Three cysteine/histidine-rich clusters which may form zinc finger-like structures are underlined. The region which is likely to adopt a coiled coil structure is also boxed [the boundaries correspond to the limit of the predicted α-helical protein segment (Gascuel and Golmard, 1988)], and within it, hydrophobic amino acids occurring at the first and fourth position of the heptad repeat are circled and underlined, respectively.

Detection of MylRAR-A (PMLRAR-A) and -B proteins

The cloned MylRAR-A and MylRAR-B cDNAs were expressed in Cos-1 cells and the proteins were revealed by Western blotting with a polyclonal antibody directed against the F region of RAR-α. MylRAR-A was detected as a doublet with apparent mol. wts of ~110 and ~120 kDa (Figure 2c, lane 5), whereas MylRAR-A and Myl2RAR-A cDNAs yielded polypeptides migrating with an apparent mol. wt of ~110 kDa (Figure 2c, lane 4, and data not shown). Whether the ~110 kDa MylRAR-A species corresponds to initiation from the second AUG (see above and Figure 1) is unknown. A polypeptide migrating at the same level as

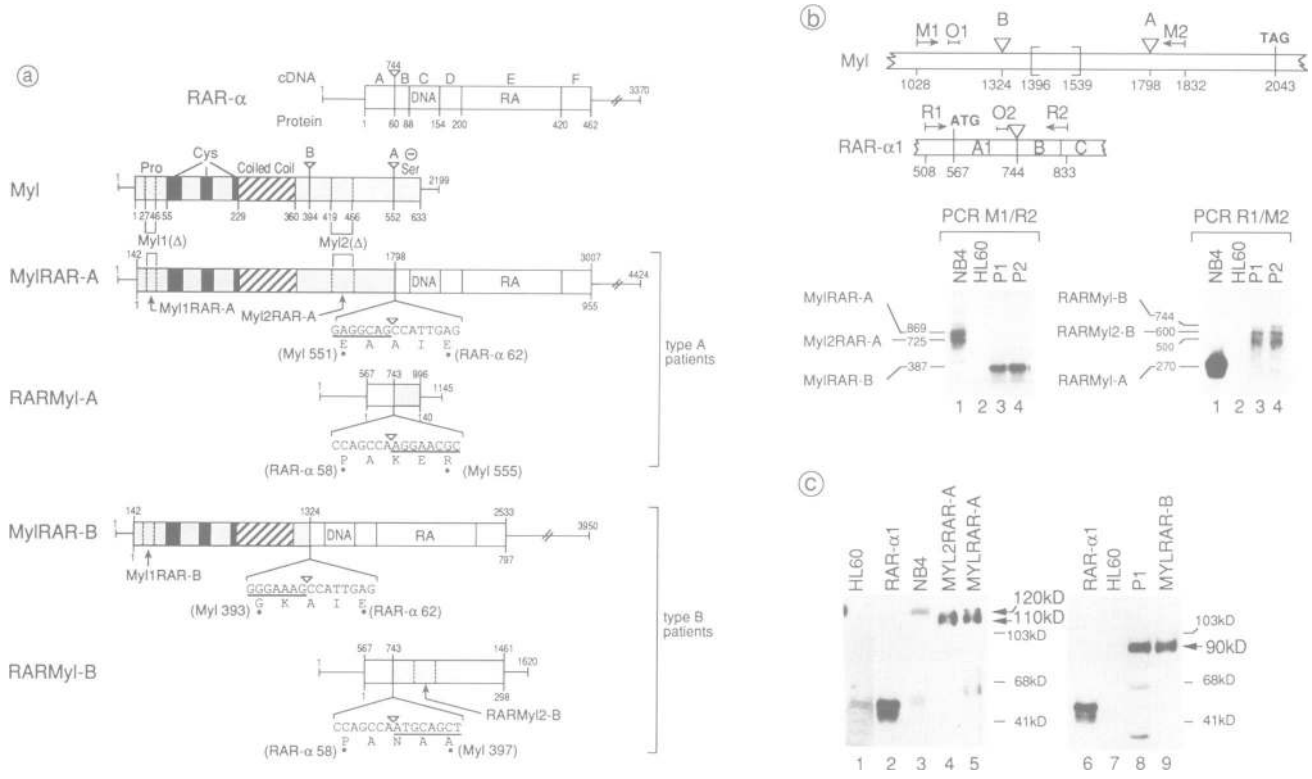

Fig. 2. Two classes of MylRAR and RARMyl transcripts and proteins. (**a**) Schematic structure of MylRAR and RARMyl cDNAs and proteins found in two classes of APL patients. On top, RAR-α1 and Myl cDNAs and proteins are schematically represented. RAR-α1 A−F regions have been defined on the basis of sequence comparisons between the different RAR subtypes (Zelent *et al.*, 1989; Kastner *et al.*, 1991). cDNA nucleotides (top lines; numbering of RAR-α1 is according to Brand *et al.*, 1990), and amino acids (lower lines) are numbered. The triangle at position 744 in RAR-α1 indicates the fusion point in MylRAR and RARMyl transcripts. A and B triangles in Myl indicate the points of fusion between Myl and RAR-α1 sequences in class A and B patients, respectively. The black boxes correspond to the three cysteine-rich clusters, which may form Zn finger-like structures in Myl. The Myl region which is likely to adopt a coiled coil structure is represented by a hatched box. The two regions of Myl which are excluded in Myl1 and Myl2 transcripts (amino acids 27−46 and 419−466, respectively) are bracketed by dashed lines. Class A and B MylRAR and RARMyl cDNAs and proteins characterized in APL class A and B, are represented below. In each case the nucleotide and derived amino acids sequence occurring at the junction is given with the Myl sequence underlined. (**b**) PCR detection of MylRAR and RARMyl fusion transcripts. The experiment is schematically represented on top. Myl and RAR-α1 cDNAs around the fusion points are depicted and the nucleotides positions corresponding to the 5′ end of the PCR primer are indicated. The presumptive Myl exon which is excluded in Myl2 by alternative splicing is bracketed. The PCR detection of MylRAR and RARMyl transcripts present in NB4 cells and in two class B APL patients (P1 and P2) is shown below in the left and right panels, respectively. PCR was performed with either the M1 and R2 primers (MylRAR amplification, left panel) or R1 and M2 primers (RARMyl amplification, right panel) on cDNA derived from NB4 cells (lane 1), HL60 cells (lane 2) and two APL class B patients (lanes 3 and 4). Amplified products were separated by electrophoresis and hybridized to end-labeled oligonucleotide probes O1 (left panel) or O2 (right panel). Amplified fragments corresponding to MylRAR-A, Myl2RAR-A, MylRAR-B, RARMyl-A, RARMyl-B and RAR2Myl-B are indicated. The ~500 bp fragment seen in the right panel in lanes 3 and 4 may correspond to an artefactual amplification product since it was not detected when other primer pairs were used. (**c**) Detection of cloned and endogenous MylRAR proteins by Western blot. 70 μg protein of whole cell extracts from either HL60 cells (lanes 1 and 7), NB4 cells (lane 3) or bone marrow cells from a class B APL patient (P1, lane 8) or 5−10 μg protein of whole cell extracts from Cos-1 cells transfected with either 5 μg of expression vectors for hRAR-α1 (lanes 2 and 6). MylRAR-A (lane 5), Myl2RAR-A (lane 4) or MylRAR-B (lane 9), have been separated by electrophoresis on a 10% SDS−acrylamide gel and analyzed by Western blotting (as described in Rochette-Egly *et al.*, 1991) with the rabbit polyclonal antibody RPα(F) directed against the F region of human RAR-α. Exposure time was 8 h for lanes 2−9 and 40 h for lane 1. The ~55 kDa species seen in lane 5 is an *in vitro* degradation product of MylRAR-A, since it was not seen when the cells were directly lysed at 100°C (not shown). Note that MylRAR-A expression vectors which do or do not contain the upstream in frame CTG (nucleotide 19) generate the same protein pattern (not shown), excluding that the 120 kD polypeptide species could be initiated at that CTG.

the ~120 kDa MylRAR-A was similarly detected using extracts from NB4 cells and cells from class A patients (Figure 2c, lane 3, and unpublished results in collaboration with Dr Pelicci's group). The same 120 kDa species could also be immunoprecipitated with an anti-Myl (amino acid 484−499) polyclonal antibody [RP(Myl)-1] and revealed by Western blotting using the anti-RAR-α antibody (data not shown), which further supports the existence of the MylRAR-A fusion protein. The polypeptide encoded by MylRAR-B cDNA migrated with an apparent mol. wt of ~90 kDa (Figure 2c, lane 9). A protein of similar size was revealed in extracts from cells of class B patients (Figure 2c, lane 8).

It is noteworthy that the level of MylRAR proteins present in APL cells was much higher than that of RAR-α [Figure 2c, lanes 3 and 8; the ~50 kDa polypeptide detected in NB4 cell extract is likely to be a MylRAR-A degradation product, since it migrated more slowly than RAR-α1 present in either Cos-1 transfected cells (lane 2) or HL60 cells (lane 1); note also that lane 1 was exposed five times longer than lanes 2−9 and that similar amounts of RAR-α RNA were found in HL60 and NB4 cells (not shown)]. In contrast, the levels of MylRAR and RAR-α transcripts appear to be very similar in both classes of APL patients (see Longo *et al.*, 1990; de Thé *et al.*, 1990; Warrel *et al.*, 1991), suggesting that MylRAR mRNAs may be more efficiently translated than

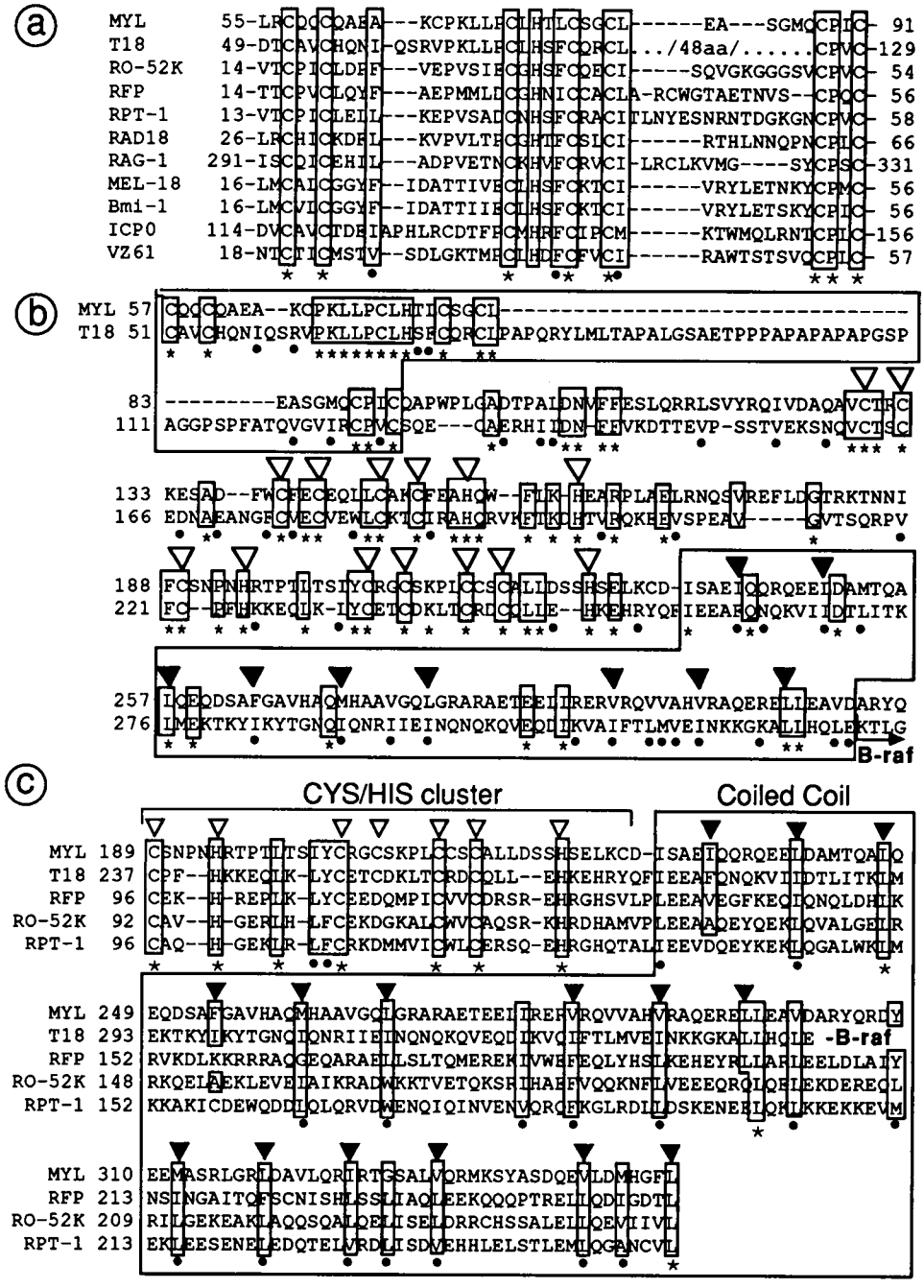

Fig. 3. Similarities between Myl and other proteins. (a) Myl belongs to a group of proteins which share a cysteine-rich motif. Residues at conserved positions are boxed. Stars indicate positions where perfect conservation occurs, whereas dots indicate positions where similar, but not identical, residues are found. (b) Alignment of Myl and T18. Since the very N-terminal sequences of Myl and T18 (Miki *et al.*, 1991) do not display significant homology, alignment is started at Cys57 and Cys51 of Myl and T18, respectively. No homology was detected also between the T18 B-Raf sequences and Myl. Residues conserved between Myl and T18 are boxed and indicated by stars. Conservative replacements are indicated by dots. The region corresponding to the cysteine-rich motif displayed in (a), as well as a region which may form a coiled coil, are boxed. Open triangles point to conserved cysteine and histidine residues belonging to the second and third Cys/His clusters in the Myl sequence. Black triangles point to the hydrophobic residues defining the heptad repeats of the putative coiled coil domain in Myl and T18. The beginning of B-Raf in the T18 sequence is indicated by an horizontal arrow. (c) Alignment of Myl, T18, RFP, RO-52K and RPT-1 in the third cysteine/histidine cluster and in the putative coiled coil. Hydrophobic amino acids defining the heptad repeat have been boxed as well as amino acids which occur at positions where identical or similar residues are found in all five proteins. Other symbols are as in Figure 3b. Note also the frequent conservation of amino acids between three or four out of the five proteins.

RAR-α mRNA and/or that MylRAR proteins are more stable than RAR-α.

Myl (PML) structural features and similarities with other proteins

The N-terminal part of Myl contains three clusters rich in cysteine and histidine residues, which may form zinc finger-like structures (underlined in Figure 1). The first of these clusters (amino acids 57–91) defines a motif which has recently been pointed out in a number of proteins [$CXXCX_{11-27}CXHX(F/L)CXXC(L/I)X_{3-48}CPXC$), see Figure 3a and Freemont *et al.*, 1991]. Several proteins of this group may be involved in cell transformation: MEL-18 is expressed in most human transformed cell lines, but not

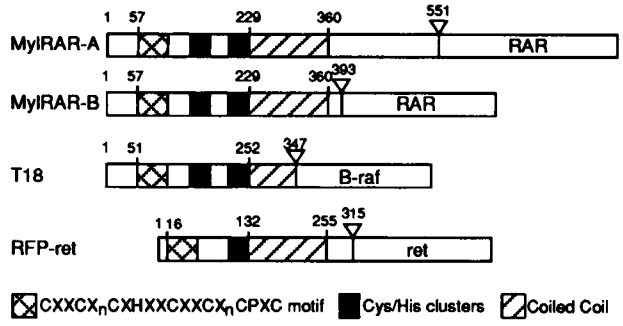

Fig. 4. Structural similarities between MylRARs and the transforming fusion proteins T18 and RFP-ret. A schematic representation of MylRAR-A, MylRAR-B, T18 and RFP-ret is displayed. Points of fusion are indicated by triangles.

in normal tissue (Tagawa *et al.*, 1990); T18 is a transforming mouse fusion protein between a novel protein containing this cysteine-rich motif and the B-*Raf* proto-oncogene [Miki *et al.*, 1991; for convenience, the (as yet not characterized) wild type protein corresponding to the N-terminal part of T18 will be hereafter referred to as T18WT]. Human RFP is fused with the *ret* proto-oncogene in a transforming protein resulting from a chromosomal translocation (Takahashi *et al.*, 1988) and mouse Bmi-1 (a close relative of MEL-18) cooperates with Myc in lymphoma development (Haupt *et al.*, 1991; van Lohuizen *et al.*, 1991a). Other proteins containing this cysteine-rich motif are involved in control of gene expression: RPT-1, which affects the expression of the IL-2 receptor (Patarca *et al.*, 1988), the herpes simplex virus immediates early gene product ICPO (a regulator of HSV gene expression) (Gelman and Silverstein, 1987), and the varicella zoster virus VZ61 protein (Davidson and Scott, 1986) which acts negatively on the expression of several varicella virus and cellular genes (Nagpal and Ostrove, 1991). This group of proteins includes also Rad18 which is a yeast protein required for repair of UV-damaged DNA (Jones *et al.*, 1988). RAG-1 which is encoded by the V(D)J recombination activating gene (Schatz *et al.*, 1989), the 52 kDa component of the RO/SSA (RO-52K) ribonucleoprotein particle which is an autoantigen in lupus erythematosus and Sjödren syndrome (Ben-Chetrit *et al.*, 1988; Chan *et al.*, 1991; Itoh *et al.*, 1991), the products of the baculovirus genes CG30, PE38, the trypanosome L/R proteins (Haupt *et al.*, 1991 for refs) and the proteins encoded by the *Drosophila Posterior Sex Comb* (psc) and *Suppressor-2 of Zeste* [Su(Z)2] genes (Brunk *et al.*, 1991; van Lohuizen *et al.*, 1991b).

The highest similarity is found between Myl and TI8, since it extends up to the fusion boundary with the B-Raf sequence which constitutes the C-terminal part of the T18 protein (Miki *et al.*, 1991) (Figure 3b). All the cysteines and histidines of the second and third Cys/His clusters are conserved between Myl and T18 with the exception of Cys213, indicating the functional importance of these residues which may be required for coordinating metal ions in zinc finger-like structures (see Berg, 1990; Vallee *et al.*, 1991 for reviews). Myl and T18 also display similarities in their third cysteine-rich cluster with RO-52K, RFP and RPT-1, where most of the cysteine and histidine residues are conserved (Figure 3c). (Note that RFP, RO-52K and RPT-1 contain only two cysteine/histidine clusters.)

Immediately C-terminal to the Myl cysteine clusters is a region (amino acids 229–360, boxed in Figure 1) which is predicted to be mostly α-helical (Garnier *et al.*, 1978; Gascuel and Golmard, 1988). This region was compared with the data bank and most of the similarities were found in protein regions which are known to form a coiled coil structure (Cohen and Parry, 1986). These proteins included myosins, keratins, dystrophin, α-actinin, flagellin, kinesin heavy chain, neurofilament triplet L protein, the *Drosophila* glued protein, tropomyosin, laminin-B2, spectrin, and Fos, Fra-1 and Fra-2 in their leucine zippers. This region of Myl contains stretches in which hydrophobic amino acids occur at every seventh position (circled in Figure 1), with the frequent presence of a hydrophobic amino acid at the fourth position (underlined in Figure 1). Interestingly, these heptad repeats are conserved between Myl, T18, RFP, RO-52K and RPT-1 (Figure 3b and c). We note also that the region which is immediately C-terminal to the cysteine cluster in RPT-1 has been reported to have a high probability of forming a coiled coil (Lupas *et al.*, 1991). Interestingly, the spacing between the different stretches of heptad repeats has also been conserved between the five proteins (Figure 3c). Therefore, Myl, T18WT, RFP, RO-52K and RPT-1 define a novel protein family. Note that RFP, RO-52K and RPT-1 are more related to each other (see Chan *et al.*, 1991 for an alignment) than to Myl or T18WT. Since T18, RFP-Ret and possibly MylRAR are transforming fusion proteins, these proteins may represent a new family of potentially oncogenic proteins. Interestingly, MylRARs, T18 and RFP-Ret retain the cysteine/histidine rich motifs and either the entirely (MylRARs and RFP-ret) or a large part of the coiled coil (possibly T18; note that the T18WT sequence C-terminal of the fusion point is not known) (see Figure 4). These structural similarities suggest that the presence of these domains may be important for the transforming potential of the fused proteins.

Other interesting structural features of the Myl protein correspond to the presence of a proline-rich N-terminus (of which a portion is deleted in the Myl1 isoform, see Figures 1 and 2a) and of a C-terminus of marked acidic character and rich in serines, among which several are potential targets for phosphorylation by casein kinase II (note that this C-terminal domain is not present in the MylRAR fusions).

Localization of Myl, MylRAR and RARMyl

The intracellular localization of Myl, MylRAR and RARMyl was analyzed by immunofluorescence performed on Cos-1 cells transfected with the corresponding expression vectors. We also used Myl(F), an epitope-tagged Myl, which consists of Myl to which the F region of the estrogen receptor (ER) is C-terminally fused (the ER F region does not possess any nuclear targeting properties, unpublished results from our laboratory). Both Myl and Myl(F) were mostly nuclear, but some staining was also seen in the cytoplasm of ~80% of the transfected cells. Characteristically both proteins were excluded from the nucleolus and concentrated in discrete speckles within the nucleus (Figure 5; panels 2 and 3). Myl(F) bearing a mutation in the cysteine-rich motif of Myl (Gln59 Cys60 → Glu59 Leu60; Myl(F)m in Figure 5, panel 7), as well as N-terminally truncated Myl and Myl(F), starting at Met312, were localized exclusively within the nucleus in most transfected cells, had lost their speckled pattern, and were more uniformly distributed (Figure 5,

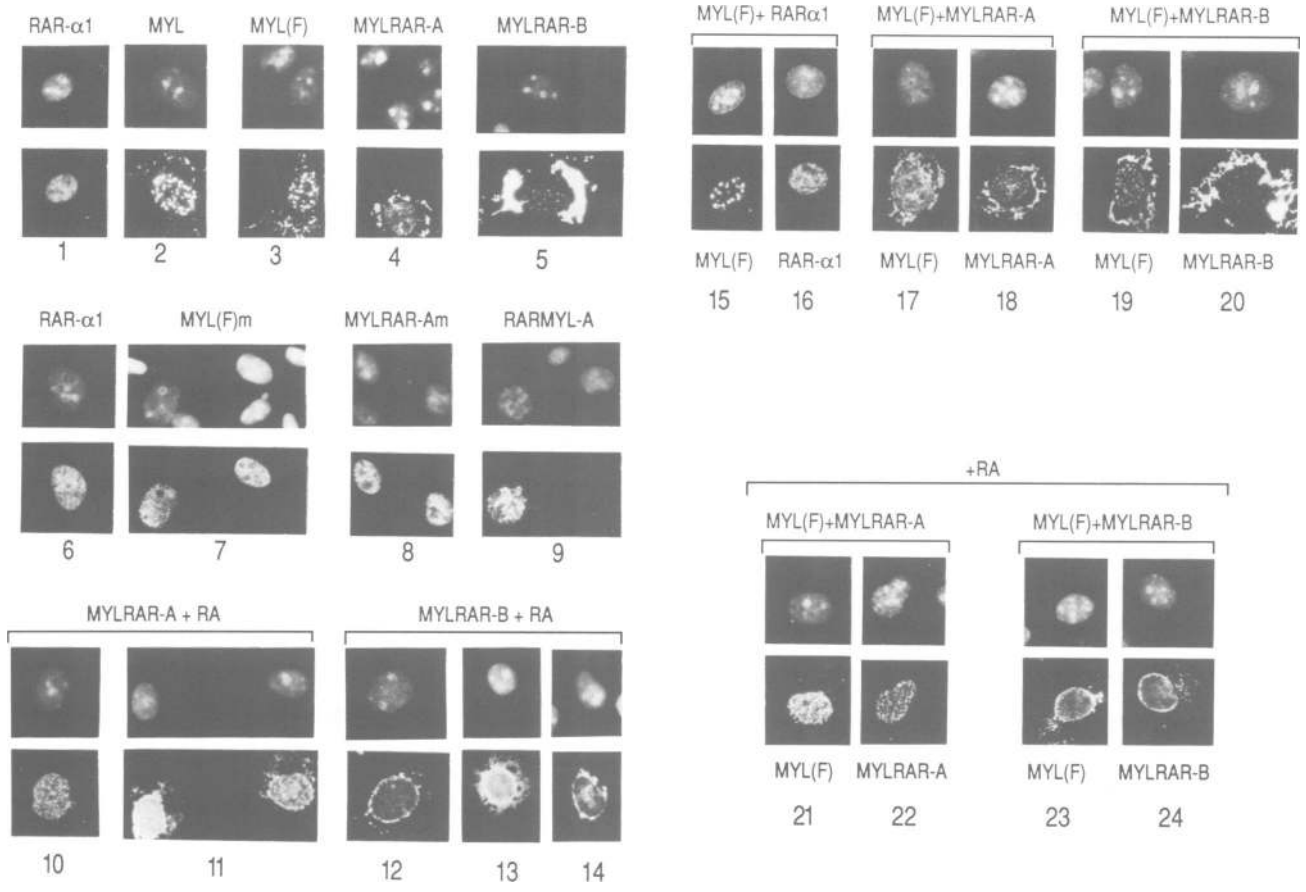

Fig. 5. Localization of RAR-α, Myl, MylRAR and RARMyl. Expression vectors encoding either RAR-α1 (panels 1, 6, 15 and 16), Myl (panel 2), the epitope-tagged Myl(F) (panels 3, 15–24), MylRAR-A (panels 4, 10, 11, 17, 18, 21 and 22), MylRAR-B (panels 5, 12, 13, 14, 19, 20, 23 and 24), Myl(F)m (panel 7), MylRAR-A(m) (panel 8) or RARMyl-A (panel 9) were transfected in Cos-1 cells, and the corresponding proteins were revealed by immunofluorescence after 24 h. In all cases the upper panel shows nuclei of transfected and untransfected cells revealed by Hoechst DNA staining and the lower panel corresponds to the immunodetection. In panels 1–14, cells have been transfected with 1 μg of expression vector, as indicated. No qualitative differences in the intracellular distribution pattern was observed when cells were transfected with lower amounts of vector (1–100 ng). In panels 15–24, cells have been transfected with 100 ng of Myl(F) and 1 μg of either RAR-α, MylRAR-A or MylRAR-B, as indicated.

panel 7, and data not shown). It is likely that Myl contains a nuclear localization signal (NLS) (Silver, 1991) located in the C-terminal portion of the protein. The KRKCSQTQCPRKVIK basic sequence (amino acids 476–490) is a possible candidate. From the uniform nuclear distribution of the Myl(F)m molecule, we conclude that the cysteine-rich motif is required for the speckled distibution of Myl.

In contrast to RAR-α localization which was exclusively nuclear and finely dispersed in all of the transfected cells (Figure 5, panels 1 and 6). MylRAR-A and MylRAR-B staining was localized both in the cytoplasm and in the nucleus, although occasionally it was only nuclear (in 10–20% of the cells in the case of MylRAR-A and in <5% in the case of MylRAR-B) (Figure 5, panels 4 and 5). The granular MylRAR nuclear distribution was clearly different from that of Myl. MylRAR-A bearing a mutation in the Myl cysteine-rich region (Gln59 Cys60 → Glu59 Leu60; MylRAR-Am in Figure 5, panel 8), as well as N-terminally truncated MylRAR-A starting at Met312 (not shown), were exclusively nuclear and more uniformly distributed, indicating that the Myl N-terminal region may modulate the nuclear translocation of MylRAR-A and its granular distribution. Interestingly, the addition of RA resulted in an increased and more uniform nuclear localization of MylRAR-A and MylRAR-B accompanied by a perinuclear accumulation of the proteins (Figure 5, panels 10–14). Note that, both in the absence and presence of RA, MylRAR-A was 'more nuclear' than MylRAR-B, possibly reflecting the presence in MylRAR-A of both Myl and RAR NLS. RA had no effect on RAR-α localization (not shown).

The localization of the 'reciprocal' protein RARMyl-A was analyzed with an antibody directed against the A1 region of RAR-α. In all of the transfected cells. RARMyl-A was essentially nuclear with a granular distribution (Figure 5, panel 9). Given its small size, RARMyl-A is likely to diffuse freely into the nucleus where it may be retained by binding to nuclear components.

Functional differences between MylRARs and RAR-α

Transcriptional activation by the different MylRAR proteins was studied in either HeLa cells or Cos-1 cells transiently cotransfected (in the presence or absence of RA) with a reporter plasmid containing the chloramphenicol acetyl transferase (CAT) gene under the control of a RA-responsive promoter, and an expression vector encoding either human RAR-α1 (hRAR-α1), hRAR-α deleted for region A [hRAR(Δ)] or a MylRAR fusion. The parental expression

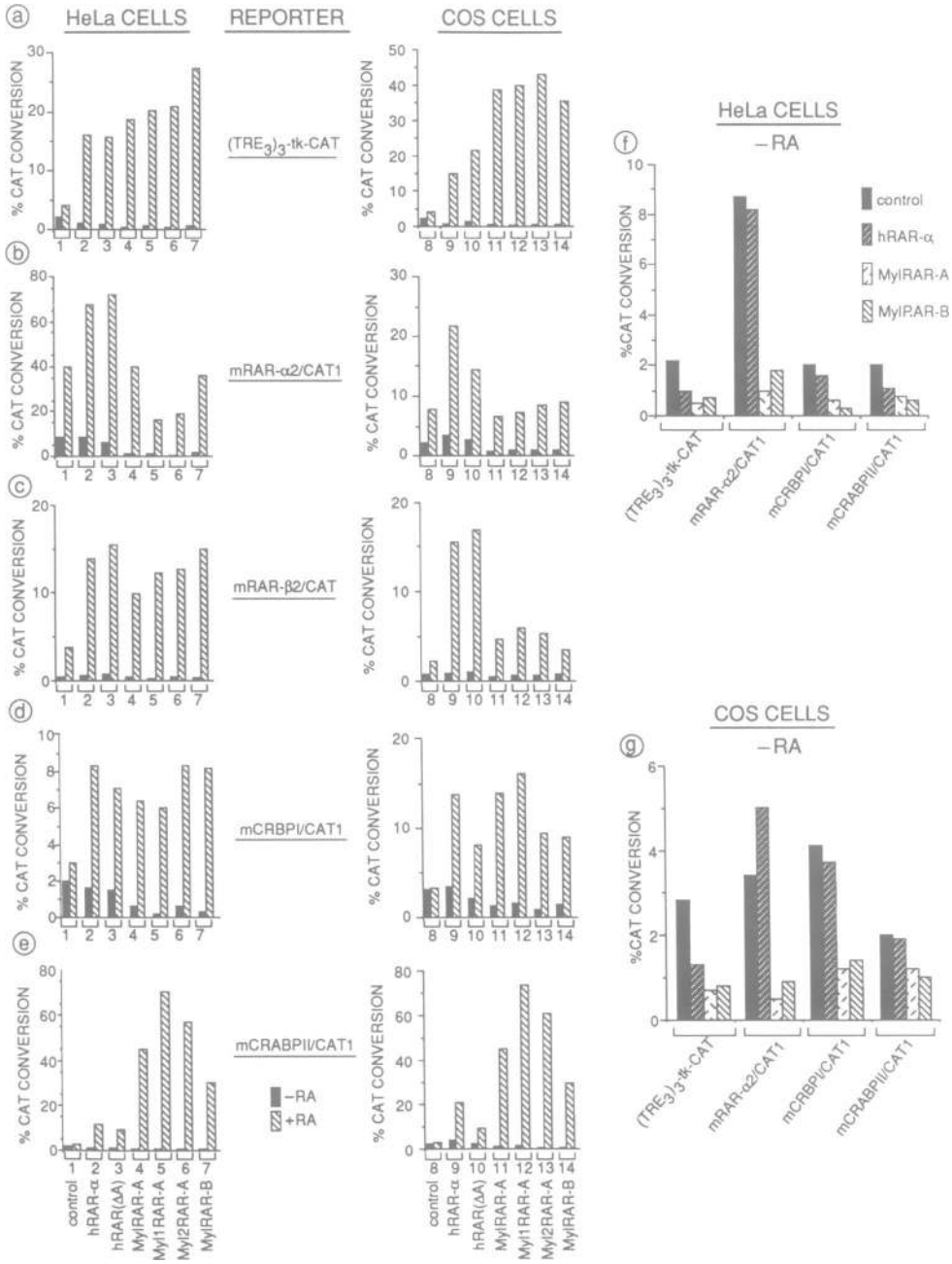

Fig. 6. Transcriptional properties of MylRARs on RA target genes. (a)−(e): HeLa cells (left panels) or Cos-1 cells (right panels) were transfected with pTL2 (lanes 1 and 8), hRAR-α1 (Brand *et al.*, 1988) (lanes 2 and 9), hRAR(ΔA) (lanes 3 and 10), MylRAR-A (lanes 4 and 11), Myl1RAR-A (lanes 5 and 12), Myl2RAR-A (lanes 6 and 13) and MylRAR-B (lanes 7 and 14), together with the following RA-responsive promoter/CAT reporter genes. (**a**) synthetic $(TRE_3)_3$-tk-CAT (Zelent *et al.*, 1989), (**b**) mouse RAR-α2 promoter [mRAR-α2/CAT1 (Leroy *et al.*, 1991)], (**c**) mouse RAR-β2 promoter [mRAR-β2/CAT (Mendelsohn *et al.*, 1991; Smith *et al.*, 1991)], (**d**) mouse cellular retinol binding protein I promoter [mCRBPI/CAT1 (Smith *et al.*, 1991)], (**e**) mouse cellular retinoic acid binding protein II promoter (mCRABPII/CAT1, B.Durand and P.Chambon, in preparation). Black bars represent basal levels in the absence of RA, and hatched bars induced levels in the presence of 10^{-6} M RA. (**f**) and (**g**): repression of the basal promoter activity of different reporter genes by MylRAR-A and MylRAR-B in HeLa and Cos-1 cells, respectively. The bars represent the CAT activity observed in the absence of RA when either mRAR-α2/CAT1, $(TRE_3)_3$-TK-CAT, mCRBPI/CAT1, mCRABPII/CAT1 were transfected with either pTL2 (control), hRAR-α1, MylRAR-A or MylRAR-B, as indicated.

vector pTL2 (Green *et al.*, 1988, and Material and methods) was used in control experiments in which RA-induced expression was due to RARs endogenous to HeLa and Cos-1 cells. The reporter plasmids were either $(TRE_3)_3$-tk-CAT (Zelent *et al.*, 1989) which contains a synthetic RARE, mRAR-α2/CAT1 (Leroy *et al.*, 1991), mRAR-β2/CAT (Smith *et al.*, 1991; Mendelsohn *et al.*, 1991), mCRBPI/CAT1 (Smith *et al.*, 1991) and mCRABPII/CAT1 (B.Durand and P.Chambon in preparation) which correspond to natural RA responsive promoters.

Cotransfections of either MylRAR-A, Myl1RAR-A, Myl2RAR-A or MylRAR-B had comparable effects on a given reporter gene, both in the absence or in the presence of RA. However, the magnitude of the response varied widely, depending on both the reporter gene promoter and the transfected cell type. In the presence of RA (Figure

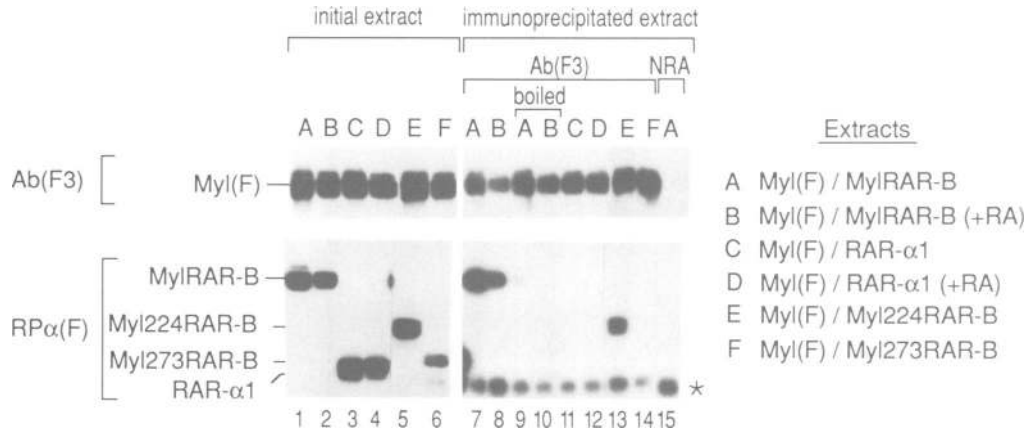

Fig. 7. Coimmunoprecipitation of Myl and MylRAR-B. Extracts A−F are from Cos-1 cells that have cotransfected with Myl(F) (Myl tagged with the ER F region) and either MylRAR-B (extracts A and B), hRAR-α1 (extracts C and D), Myl224RAR-B (extract E) or Myl273RAR-B (extract F). Cells used for preparation of extracts B and D have been grown in the presence of 10^{-6} M RA. Extracts were immunoprecipitated with the anti-ER F region Ab(F3) monoclonal antibody or a non-reactive ascite fluid (NRA), as indicated. Immunoprecipitates were loaded on two SDS gels which were revealed with Ab(F3) (upper panel) or the anti-RAR-α F region polyclonal antibody RPα(F) (lower panel). 5 μl of each cell extract (1/200 of the material used for immunoprecipitation) was loaded in lanes 1−6. Extracts in lanes 9 and 10 have been boiled for 15 min prior to immunoprecipitation. Starts point to immunoglobulins present in the immunoprecipitated samples.

6a−e), $(TRE_3)_3$-tk-CAT and mCRBPI/CAT1 expressions were stimulated to similar extent by hRAR-α1 and MylRARs, whereas MylRARs were less efficient than hRAR-α1 at stimulating the mRAR-α2/CAT reporter. In contrast, MylRARs (particularly MylRAR-A) were more efficient than hRAR-α1 at activating mCRABPII/CAT1 expression. In the case of mRAR-β2/CAT, hRAR-α1 and MylRARs activated similarly in HeLa cells, whereas hRAR-α1 was more efficient than MylRARs in Cos-1 cells. The apparent affinity of MylRARs for RA was similar to that of hRAR-α1, with in both cases ED_{50} of $\sim 5 \times 10^{-9}$ M for the stimulation of the CRABPII promoter (data not shown).

Interestingly, in the absence of RA, MylRARs had a repressive effect on the basal activity of certain reporter gene promoters, under conditions where hRAR-α1 cotransfection resulted in no or much less repression (Figure 6f−g). This differential repression was particularly clear in the case of the mRAR-α2 and mCRBPI reporter gene promoters (Figure 6f and g). Note that in all cases hRAR-α1 and hRAR(ΔA) had similar effects, which indicates that the promoter- and cell-specific differences between hRAR-α1 and MylRAR activation and repression functions are due to the presence of the Myl sequence in MylRARs.

Since the reciprocal translocation product, RARMyl, contains a Myl region rich in acidic amino acids (see above and Figures 1 and 2a), we tested whether RARMyl-A could squelch the activity of two acidic activators in cotransfection experiments similar to those reported by Tasset *et al.* (1990). No effect was found on Gal4 or Gal-VP16 activated transcription, nor could RARMyl-A interfere with hRAR-α1-mediated gene activation (data not shown).

Association between Myl and MylRAR

Since Myl contains a putative coiled coil structure (see above), we tested wheter Myl derivatives could associate. Cos-1 cells were transiently cotransfected either with vectors expressing the epitope-tagged Myl(F) and MylRAR-B, or Myl(F) and hRAR-α1 vectors. Extracts were prepared from both sets of transfected cells and subjected to immunoprecipitation with either the Ab(F3) monoclonal antibody specific for the ER(F) tag, or a control ascitic fluid. Immunoprecipitates were then separated by electrophoresis on two identical gels which were analyzed by Western blotting with Ab(F3) and an anti-RARα-F region polyclonal antibody [RPα(F)]. Myl(F) and MylRAR-B, but not Myl(F) and hRAR-α1, coimmunoprecipitated (Figure 7, lanes 7−8 and 11−12). A similar coimmunoprecipitation was observed when the cells were grown in the presence of RA (Figure 7, lane 8). This coimmunoprecipitation was not observed when the cell extract was boiled prior to immunoprecipitation (Figure 7, lanes 9 and 10) nor was it observed between the ER and MylRAR-B (not shown), excluding potential antibody artefact. Similar results were obtained with *in vitro* cotranslated MylRAR-A and Myl(F) (data not shown). Thus, the formation of Myl(F)/MylRAR complexes suggests that Myl and MylRAR may actually form homo- and heterodimers in solution. We tested the ability of two N-terminally truncated MylRAR-B molecules to associate with Myl(F). Myl224RAR-B is missing all sequences N-terminal to Glu224 (which include the cysteine clusters) and Myl273RAR-B is missing all sequences N-terminal to Ala273 (missing the cysteine clusters and part of the putative coiled coil). Myl224RAR-B, but not Myl273RAR-B, coimmunoprecipitated with Myl(F) (Figure 7, lanes 13 and 14). Thus the conserved cysteine-rich regions are not involved in Myl/MylRAR association, whereas the coiled coil region is required, most probably by providing a surface for dimerization. In a similar experiment, no coimmunoprecipitation could be observed betwen MylRARs and an epitope-tagged RAR-α (not shown), indicating that, although a putative dimerization interface has been proposed to exist in RAR-α (Forman and Samuels, 1990), this domain does not allow formation of stable RAR-α/MylRAR dimers in solution.

To show the Myl/MylRAR complexes are formed within cells, Cos-1 cells were cotransfected with Myl(F) and MylRAR, or Myl(F) and hRAR-α1, and the intracellular localization of Myl(F) was investigated. When cotransfected with hRAR-α1, Myl(F) exhibited the expected Myl pattern

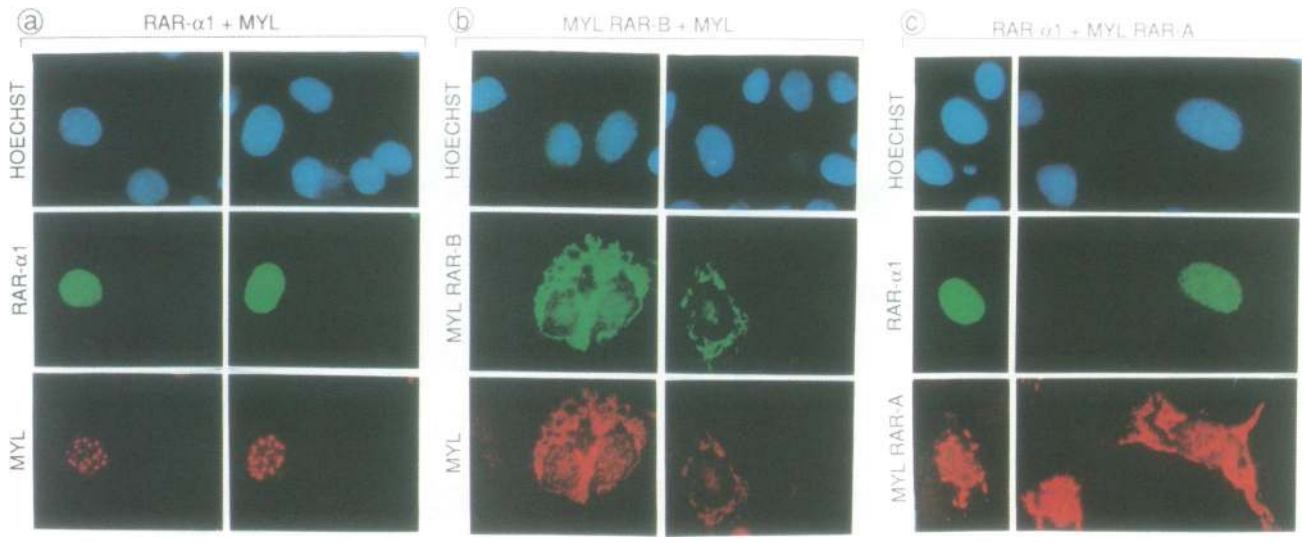

Fig. 8. Colocalization of Myl and MylRAR-B, Cos-1 cells cotransfected with Myl (1 μg) and hRAR-α1 (100 ng) (**a**), Myl (100 ng) and MylRAR-B (1 μg) (**b**) and MylRAR-A (1 μg) and hRAR-α1 (100 ng) (**c**) were subjected to double immunofluorescence with the anti Myl RP(Myl)-1 polyclonal antibody, detecting Myl and MylRAR-A, but not MylRAR-B, the Ab(9α)F monoclonal antibody detecting hRAR-α1 and MylRAR-B in (a) and (b), and Ab10α1(A1), detecting hRAR-α1 in (c). Top panels shows the DNA Hoechst staining, middle panels the RARα or MylRAR-B localization and lower panels the Myl or MylRAR-A localization, as indicated.

(Figure 5, panel 15). In contrast, when cotransfected with MylRAR-A or -B, Myl(F) displayed the typical pattern of MylRAR both in the absence (compare panels 17 and 19 with panels 18 and 20 in Figure 5) or presence of RA (compare panels 21 and 23 with panels 22 and 24 in Figure 5). To support further the colocalization of coexpressed Myl and MylRAR, cells were transfected with either Myl and MylRAR-B or Myl and hRAR-α1 and the distribution of Myl, MylRAR-B or hRAR-α1 were revealed in the same cells by double labelling immunofluorescence. As expected, coexpression of Myl and hRAR-α1 did not perturb either Myl or hRAR-α normal localizations (Figure 8a). In contrast, when coexpressed with MylRAR-B, Myl adopted the exact MylRAR-B localization pattern (Figure 8b). Therefore, Myl is likely to be associated with MylRARs within the cell, and this association appears to interfere with its 'normal' intracellular localization. In similar double labelling immunofluorescence experiments, transfection of MylRAR-A in excess [revealed specifically with RP(Myl)-1] together with hRAR-α1 [revealed with Ab10α1(A1), directed against the RAR-α1 A1 region] did not alter significantly the nuclear localization of RAR-α1 (Figure 8c), further indicating that MylRARs and RAR-α1 do not associate.

Discussion

We report here the molecular characterization of two classes (A and B) of reciprocal RNA transcripts and proteins which result from the chromosomal t(15;17) translocation specific to APL. In both cases the breakpoint in the RAR-α gene is located in the intron separating the exon encoding region B from more upstream regions which contain the two RAR-α gene promoters and the multiple exons encoding the different 5′-UTRs and A regions of the various RAR-α isoforms (Brand *et al.*, 1990; Leroy *et al.*, 1991; P.Leroy and P.Chambon, unpublished results). In contrast, there are two classes of breakpoints in the Myl gene, resulting in MylRAR-A and B fusion transcripts and proteins, and in the corresponding RARMyl-A and B reciprocal counterparts. The existence of these two A and B classes is in agreement with previous reports, which showed that there are two different size classes of abnormal RAR-α transcripts in APL patients (Longo *et al.*, 1990; Warrel *et al.*, 1991). Note that we detected only one class of transcripts (either A or B) in any given patient, thus confirming the clonal origin of the APL tumoral cells. The PCR assay that we used here offers a sensitive and reliable assay to discriminate between the two t(15;17) translocation classes, thus making possible an investigation into whether A and B patients respond similarly to RA treatment.

Myl contains a cysteine-rich motif which is shared by a newly identified group of proteins, of which some appear to exert regulatory functions in the nucleus. The occurrence of this motif, in species as distant from human as yeast and trypanosome (see Haupt *et al.*, 1991 for refs), suggests that it could perform an evolutionarily conserved function. It may possibly form a new zinc-coordinated finger structure (Berg, 1990; Vallee *et al.*, 1991), which could correspond to a novel class of DNA (or RNA) binding motif. Two additional cysteine-rich motifs are present in the N-terminal region of Myl which also appear to be evolutionarily conserved (see Figure 3). Immediately C-terminal to these cysteine-rich motifs, Myl contains a region which is likely to form a coiled coil. In agreement with the existence of this structure, we have shown that Myl can form a complex with MylRAR (presumably as a heterodimer) and that the Myl coiled coil is required for this association. In addition, Myl contains a proline-rich region, as well as an acidic region, both of which have been found in the activating domains of some transcription factors (Mitchell and Tjian, 1989). The overall N-terminal structure (putative Zn fingers followed by a coiled coil, see Figure 3) has been conserved between Myl, T18WT, RFP, the putative transcription factor RPT-1 and the 52 kDa component of the RO/SSA ribonucleoprotein particle. These proteins define a novel protein family, and they may therefore perform related functions. Since three out of these five proteins (Myl, T18WT and RFP) have been characterized in the context of oncogenic fusion proteins,

these proteins represent a new class of proteins which may be involved in cell transformation.

Could Myl be a transcription factor? This idea is supported by its similarities with other proteins (e.g. ICPO) and RPT-1) which possess regulatory functions and could be transcription factors, and by the presence of both putative DNA binding and activating domains within the Myl sequence. The possible presence of an activation function in the Myl portion present in MylRARs is supported by the strong activation of the CRABPII promoter by MylRARs, but neither by RAR-α nor by RAR-α deleted for the A region. Note, however, that we did not observe any stimulation of Gal4 reporter genes transfected into Cos-1 cells together with a vector expressing the entire Myl protein fused to the Gal4 DNA binding domain (Tora *et al.*, 1989; our unpublished results). On the other hand, the speckled nuclear distribution of Myl as well as its similarity with the RO-52K protein may be more consistent with a role of Myl as a component of ribonucleoprotein complexes. In this respect, it is noteworthy that (i) two proteins of the same group, ICPO and the RO/SSA 52 kDa ribonucleoprotein, have been reported to display speckled nuclear localizations (Gelman and Silverstein, 1987; Ben-Chetrit *et al.*, 1988); (ii) the mutation of two residues in the conserved cysteine-rich motif is sufficient for abolishing the speckled appearance and making Myl fully nuclear. Spliceosomes have also been shown to exhibit a speckled nuclear distribution (Fu and Maniatis, 1990; Carmo-Fonseca *et al.*, 1991; Gall, 1991). However, we note that the spliceosome-specific antibody SC35 (Fu and Maniatis, 1990) does not label the Myl speckles (our unpublished results).

Irrespective of the actual function of Myl, there is little doubt that its fusion products with RAR-α are responsible for the lack of differentiation of promyelocytes, which is tightly correlated with the t(15;17) translocation and the expression of abnormal RAR transcripts in APL (Longo *et al.*, 1990; Warrel *et al.*, 1991). In this respect it is striking that the portion of Myl which is fused with RAR in MylRARs contains a cysteine-rich and coiled coil motifs which are also found in the N-terminal moiety of the two oncogenic fusion proteins T18 and RFP-ret (see above). Although it cannot be excluded that the reciprocal RARMyl fusion products could be partially responsible for the block of promyelocyte differentiation in APLs, we will restrict below our discussion to the possible role of MylRARs in this block since it is relieved by RA treatment. Any model aimed at explaining the effects of MylRARs on promyelocyte differentiation must ultimately account for two facts: (i) normal Myl and RAR-α are apparently synthesized by the non-translocated allele genes in APL cells; (ii) supra-physiological RA concentrations are required in the blood of APL patients to result in remission (Warrel *et al.*, 1991). If one assumes that Myl and/or RAR are involved in the control of normal myelocyte differentiation, it follows that MylRARs must not only be unable to perform the functions of Myl and RAR-α, but also act as *trans*-dominant inhibitors of these functions; furthermore increased concentrations of RA would be required to relieve this negative *trans*-dominance.

Two mechanisms can be considered. Firstly, MylRARs could interfere with a normal physiological control of expression of RA-responsive genes, which may be crucial for differentiation of promyelocytes to granulocytes. Note in this respect that RA is known to induce the differentiation of several non-acute promyelocytic leukemia cell lines (Lübbert and Koeffler, 1988). MylRAR appears to be present at much higher levels than wild type RAR-α in APL cells (see Figure 2c), and therefore its effects are likely to be dominant to those of RAR-α. We have shown, as well as others (de Thé *et al.*, 1991; Kakizuka *et al.*, 1991; Figure 6) that MylRAR-A and MylRAR-B have similar transactivating properties which can differ markedly from those of RAR-α1. These differences between MylRARs and RARs may be due to the presence in MylRAR of the Myl dimerization interface and probable MylRAR homodimers may have different DNA binding/transactivating properties than RAR-α, which does not homodimerize efficiently and requires an additional nuclear factor to bind to RAREs (Leid *et al.*, 1992). Interestingly, MylRARs can repress the basal activity of some RA target gene promoters in the absence of RA, whereas the same promoters are transactivated in the presence of RA (Figure 6). It is possible that the actual level of RA is too low in the blood of APL patients to activate MylRARs which would then act as repressors. Increasing the RA concentration in APL patients could be necessary to convert MylRARs into RAR-like activators, and also to improve their transfer into the nucleus, as suggested by the results shown in Figure 5. In other words, at RA physiological concentrations, the repressive effect of MylRARs would dominate, and supra-physiological RA concentrations would be necessary to convert MylRARs into activators of RA target genes. In this respect, we note that the basal activity of the RAR-α2 promoter is particularly sensitive to MylRAR repression. Whether expression of this RA-inducible RAR-α isoform is required for the expression of RA target genes possibly involved in promyelocyte differentiation, remains to be seen.

On the other hand, RA may not be required for the physiological differentiation of promyelocytes, which would be dependent on Myl, whose function would be *trans*-dominantly repressed by MylRARs in APL. This negative effect could result from the formation of inactive and inappropriately located Myl/MylRAR complexes (possibly heterodimers) unable to perform the normal function of Myl dimers. If Myl is a transcription factor, Myl/MylRAR heterodimers might be unable to transactivate Myl target genes because a Myl activation function (perhaps corresponding to the acidic Myl C-terminal region) may be lacking in MylRARs. This defect would then be compensated when the activation function in the RA binding domain present in MylRARs is activated by RA.

In conclusion, although we have characterized the products resulting from the APL translocation, the present study of their functional properties does not lead to a clearcut model accounting for the remarkable efficiency of RA in the treatment of APL patients. It cannot be excluded also that the fusion products by Myl and RAR could block myelocyte differentiation of modulating the activity of genes whose expression is not normally controlled by Myl and/or RAR-α. Elucidating the physiological function of Myl is an obvious prerequisite to further progress in the understanding of the molecular mechanisms leading to altered myelocyte differentiation in APL. That Myl may perform important functions is strongly suggested by its structural similarities with a new group of proteins, of which some have oncogenic properties when fused with other proteins.

Three studies reporting the characterization of Myl and MylRAR fusion transcripts have been published since

completion of the present report (de Thé *et al.*, 1991; Kakizuka *et al.*, 1991; Pandolfi *et al.*, 1991).The fusion transcript identified by both Pandolfi *et al.* (1991) and de Thé *et al.* (1991) is identical to MylRAR-A, whereas Kakizuka *et al.* (1991) have isolated a fusion transcript identical to MylRAR-B. It is interesting to note that our Myl sequence is different from those reported by all three groups, which are themselves different from each other. In all cases the reported Myl sequences are identical at least up to the point of fusion with RAR-α sequence in MylRAR-A (Myl amino acid 552). In the case of Kakizuka *et al.* (1991), the divergence which starts at amino acid 553 is due to an 8 bp insertion which may correspond to a mini-intron. The divergence with the de Thé *et al.* (1991) and Pandolfi *et al.* (1991) Myl sequence (which are different from each other) occurs at nucleotide 1851 (amino acid 570) and is likely to correspond to a splice junction. The significance of these differences which may reflect a complex pattern of alternative splicing in the C-terminal region of Myl is unknown.

Materials and methods

Cloning of cDNAs and PCR detection of MylRAR and RARmyl

Myl cDNAs were cloned as follows. A three step anchored PCR walk was performed on NB4 cDNA to isolate Myl cDNA sequences upstream of the fusion point with RAR-α. Typically, 5 μg of total NB4 RNA (Lanotte *et al.*, 1991) was reverse transcribed with a specific RAR-α region B anti-sense oligonucleotide and the resulting cDNA dG-tailed. Two rounds of anchored PCR were then performed with two nested 3′ primers, as described (Loh *et al.*, 1989; Kastner *et al.*, 1990; Zelent *et al.*, 1991). The first and second anchored PCR steps yielded sequences up to nucleotides 969 and 279, respectively (see Figure 1). The 3′ Myl sequences have been cloned by PCR using two nested Myl primers (nucleotides 1712−1731 and 1742−1762, respectively) and oligo(dT)-containing primers, as described (Frohman and Martin, 1991). All sequences have been determined on at least three independent clones. PCR analysis of MylRAR and RARMyl transcripts was carried out as follows: 5 μg total RNA was reverse-transcribed with either a RAR-α C region antisense oligonucleotide primer (MylRAR amplification) and with an oligo(dT)-containing primer (RARMyl amplification). One tenth of the reverse transcription reaction was employed for PCR. 35 cycles of PCR were performed (1 min at 94°C, 2 min at 60°C, 3 min at 72°C) with 100 pmol of each primer in a buffer containing 10 mM Tris−HCl pH 8.8, 1 mM $MgCl_2$, 50 mM KCl, 200 μg/ml BSA and 100 μM dNTPs. Primer sequences were: M1, 5′-GAGCTGCTGGAGGCTGTGGA; M2, 5′-TCTTCCGAGCTGCTGATCAC; R1, 5′-GCCCACCAGAGG-CCCCCTGC: R2, 5′-AAAGCAAGGCTTGTAGATGC.

Plasmid constructions

MylRAR-A (see Figures 1 and 2a) expression vector was constructed from three fragments produced by restriction digestion of PCR-amplified products: *Bgl*II−*Kpn*I (nucleotides 107−1056). *Kpn*I−*Sac*II (nucleotides 1057−1864) and *Sac*II−*Eco*RI (nucleotides 1864−3009). Note that *Eco*RI and *Sac*II are new sites created by PCR; the creation of *Sac*II site in the RAR B region changes the sequence CCCCGC into CCGCGG (nucleotides 470−475 in Petkovich *et al.*, 1987, encoding ProArg) and the *Eco*RI site is located just downstream of the RAR-α stop codon (TGAATTC). These fragments were ligated together into the pTL2 expression vector (a gift of Tom Lufkin) between the *Bgl*II and *Eco*RI sites. pTL2 is identical to pSG5 (Green *et al.*, 1988), but possesses a *Bgl*II−*Kpn*I−*Sac*I−*Pst*I−*Sma*I−*Not*I−*Hin*dIII−*Bam*HI−*Eco*RI polylinker as cloning sites. The resulting plasmid (MylRAR-Aϕ) was linearized with *Bgl*II and the 106 5′-terminal Myl nucleotides were inserted as a *Bgl*II fragment (obtained by PCR) generating MylRAR-A. Both MylRAR-Aϕ and MylRAR-A generated the same 120 and 110 kDa proteins with similar efficiency when transfected into COS-1 cells (not shown). Myl1RAR-A (see Figures 1 and 2a) expression vector was made by replacing the *Bgl*II−*Kpn*I fragment of MylRAR-Aϕ with the corresponding fragment of Myl1 obtained by PCR from NB4 cell RNA. Myl2RAR-A (see Figures 1 and 2a) expression vector was obtained by replacing the *Kpn*I−*Sac*II fragment of MylRAR-Aϕ with the corresponding Myl2RAR-A fragment which was obtained by PCR performed on NB4 cell RNA. MylRAR-B (see Figures 1 and 2a) expression vector was generated by replacement of the *Kpn*I−*Sac*II fragment of MylRAR-Aϕ with the corresponding fragment obtained by PCR from class B patient RNA. The Myl expression vector was made by inserting the Myl 106−2058 nucleotide sequence (see Figure 1) between the *Bgl*II and *Eco*RI sites of pTL2. Myl(F) expression vector was constructed by inserting between *Bgl*II and *Eco*RI of pTL2, the Myl 106−2040 sequence fused with the F region of the human estrogen receptor (nucleotides 1869−2020, Green *et al.*, 1986). A *Xba*I site was created at the Myl-ER junction with the nucleotide sequence being CTAATTCCTTCTAGAACTAGC, which encodes L-T-P-S-R-T-S. Note that a proline residue has been introduced to allow for flexibility between Myl and ER sequences. The Q59C60−E59L60 mutation present in Myl(F)m and MylRAR-A(m) vectors was created by site-directed mutagenesis changing CAATGC (nucleotides 316−321; see Figure 1) into GAATTC and thus creating an *Eco*RI site. RARMyl-A vector was constructed by subcloning into the *Eco*RI site of pSG5 the fragment obtained by PCR from NB4 cDNA with the primer pair 5′-ATGAATTCCACCATGGCCAGCAACAGCAGCT and 5′-ATGAATTCTTTGGGACTCAGAGACTAAA. hRAR(ΔA) expression vector was constructed by PCR amplification of the sequence encoding the B−F regions of hRAR-α, and subcloning into *Kpn*I−*Eco*RI sites of pTL2. Amplification was performed on MylRAR-A with the 5′ oligonucleotide 5′-ATGGTACCACCATGGCCATTGAGACCCAGAGCAG and the 3′ oligonucleotide 5′-ATGAATTCACGGGGAGTGGGTGGC. Myl224RAR-B and Myl273RAR-B were made by deleting in MylRAR-B nucleotides upstream of nucleotide 811 and nucleotide 958 (Figure 1), respectively. The resulting inserts are cloned between *Bgl*II and *Eco*RI sites in pTL2 and their 5′ sequences are AGATCTCCACCATGGAGCTC . . ., encoding MetGluLeu . . . (Myl224RAR-B) and AGATCTCCACCATG GCCGAG . . . encoding MetAlaGly . . . (Myl273RAR-B). All cloned fragments which have been obtained by PCR have been resequenced.

Antibodies, immunofluorescence, immunoprecipitation and Western blotting

RP(Myl)-1 was raised against the synthetic peptide CPRKVTKMEGEEGKE (underlined by dashes in Figure 1), which was coupled to ovalbumin via the cysteine residue as described (Rochette-Egly *et al.*, 1991). Immunization of rabbits and antiserum preparation was as in Gaub *et al.* (1989). The immune serum was characterized by Western blot and immunofluorescence analysis on cloned Myl, Myl(F) and MylRAR-A. Other antibodies used in this study were: Ab9α(F), a monoclonal antibody directed against the F region of RAR-α (Gaub,M.P., Rochette-Egly,C., Lutz,Y., Ali,S., Matthes,H., Schever,I. and Chambon,P., in preparation), which detects RAR-α1, MylRAR-A, MylRAR-B and MylRAR-A(m); Ab10α1(A1), a monoclonal antibody directed against the A1 region of hRAR-α1 (M.P.Gaub, in preparation) which detects hRAR-α1 and RARMyl-A; Ab(F3), a monoclonal antibody directed against the F region of the human estrogen receptor (hER) (Metzger,D., Ali,S., Lutz,Y., Bellocq,J.P., in preparation) which detects Myl(F) and Myl(F)m.

Immunofluorescence (see Figure 5) was performed as follows. Transfected Cos-1 cells, cultured in Leighton tubes (Costar) in Dulbecco's medium supplemented with 5% delipidized fetal calf serum, were processed for the immunodetection essentially according to Lutz *et al.* (1988) except that the methanol/acetone steps were omitted. Control experiments where the fixation/permeabilization procedure was replaced by a methanol/acetone procedure showed also no modification in the antigen distribution nor did the absence or presence of the detergent Triton X-100 modify the observed localization. Antibody RP(Myl)-1 was incubated overnight at a 1:500 dilution in PBS. Hybridoma culture supernatants of Ab10α(A1), Ab9α(F) and Ab(F3) were used at dilutions of 1:1, 1:10 and 1:200, respectively. The Texas Red conjugated second antibody (Jackson Immunoresearch Laboratories), was incubated for 1 h and diluted 40 times for anti-rabbit antibody and 200 times for anti-mouse antibody. For double labelling, Myl and MylRAR-A [detected with RP(Myl)-1 diluted at 1:500] were revealed with Texas Red conjugated anti-rabbit IgG second antibody (Jackson Immunoresearch Laboratories) diluted 40 times and MylRAR-B or RAR-α [detected with Ab9α(F) diluted at 1:10] were revealed with fluorescein conjugated anti-mouse IgG second antibody (Jackson Immunoresearch Laboratories) diluted 10 times.

Immunoprecipitation reactions were carried out as follows. Cos-1 cells (9 cm petri dishes) were transfected with 2 μg of each expression vector, as indicated in Figure 7. Cells were washed with PBS and scraped in 1 ml per dish of RIPA buffer (10 mM Tris−HCl, pH 7.5, 120 mM NaCl, 1% NP-40, 1% deoxycholate, 0.1% SDS, 1 mM PMSF and PIC (leupeptin, aprotinin, pepstatin, antitrypsin and chymostatin at 0.5 μl/ml each). After 15 min on ice, the lysed cells were spun at 5000 r.p.m. for 15 min at 4°C. Immunoprecipitations (1 ml extract) were performed as described (Rochette-Egly *et al.*, 1991). Immunoprecipitates were separated on 7.5% SDS−acrylamide gels, and immunoblotted with Ab(F3) or RPα(F) and labelled protein A, as described (Rochette-Egly *et al.*, 1991).

Western blotting has been performed as described (Rochette-Egly *et al.*, 1991). RPα(F) (M.P.Gaub *et al.*, in preparation) has been used at a 1:500 dilution.

Analysis of MylRAR transcriptional properties

HeLa and Cos-1 cells were cultured in Dulbecco's medium containing 5% delipidized fetal calf serum. Transfections were performed by the calcium phosphate precipitation method. For each 9 cm Petri dish, transfected DNA consisted of 1 μg of expression vector, 1 μg of internal control plasmid pSG5-LacZ, 5 μg of CAT reporter plasmid (or 2 μg in the case of $(TRE_3)_3$-tk-CAT) and Bluescript SK^- DNA to a total amount of 20 μg. After 48 h, extracts were prepared and normalized for β-galactosidase activity. CAT assays were performed with amounts of extracts corresponding to equal β-galactosidase activity. In all cases the results correspond to the average of at least two independent transfection experiments (±20%). Addition of RA resulted consistently in a reduction of β-galactosidase activity when either RARs or MylRARs were cotransfected. To exclude that the observed CAT activity was biased by an effect of RAR or MylRAR on the SV40 promoter which directs expression of the *LacZ* gene in pSG5-LacZ, an RSV−*LacZ* recombinant was also used as an internal control; similar results were obtained. Furthermore, no RA-dependent induction or repression of CAT activity was obtained when a non-RA-responsive reporter (CRABPI/CAT) was used, both with RAR and MylRAR, supporting the validity of the β-galactosidase normalization (unpublished result).

Acknowledgements

We thank Drs X.D.Fu and T.Maniatis for the anti-SC35 antibody, S.Ali and D.Metzger for the Ab(F3) antibody, C.Reibel, I.Scheuer, N.Jung and V.Schultz for technical assistance; P.Leroy, H.Nakshatri, S.Nagpal, C.Mendelsohn and T.Lufkin for plasmids; M.Saunders and M.Leid for critical reading of the manuscript; A.Staub, F.Ruffenach and I.Colas for oligonucleotide synthesis, and members of the lab for fruitful discussions. We thank also C.Werlé, B.Fournaise, B.Boulay and J.M.Lafontaine for artwork and photography, the cell culture staff for cells and the secretaries for typing the manuscript. This work was supported by funds from the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the Association pour la Recherche sur le Cancer, and the Fondation pour la Recherche Médicale, A.P. was supported by a fellowship from the Ministère des Affaires Etrangères.

References

Alcalay,M., Zangrilli,D., Pandolfi,P.P., Longo,L., Mencarelli,A., Giacomucci,A., Rocchi,M., Biondi,A., Rambaldi,A., Lo Coco,F., Diverio,D., Donti,E., Grignani,F. and Pelicci,P.G. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1977–1981.

Ben-Chetrit,E., Chan,E.K.L., Sullivan,K.F. and Tan,E.M. (1988) *J. Exp. Med.*, **167**, 1560–1571.

Berg,J.M. (1990) *J. Biol. Chem.*, **265**, 6513–6516.

Borrow,J., Goddard,A.D., Sheer,D. and Solomon,E. (1990) *Science*, **249**, 1577–1580.

Brand,N., Petkovich,M., Krust,A., Chambon,P., de Thé,H., Marchio,A. and Dejean,A. (1988) *Nature*, **332**, 850–853.

Brand,N., Petkovich,M. and Chambon,P. (1990) *Nucleic Acids Res.*, **18**, 6799–6806.

Breitman,T.R., Selonick,S.E. and Collins,S.J. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2936–2940.

Brockes,J.P. (1989) *Neuron*, **2**, 1285–1294.

Brunk,B.P., Martin,E.C. and Adler,P.N. (1991) *Nature*, **353**, 351–353.

Carmo-Fonseca,M., Pepperkok,R., Sproat,B.S., Ansorge,W., Swansson,M.S. and Lamond,A.I. (1991) *EMBO J.*, **10**, 1863–1873.

Castaigne,S., Chomienne,C., Daniel,M.T., Ballerini,P., Berger,R., Fenaux,P. and Degos,L. (1990) *Blood*, **76**, 1704–1709.

Cavener,D.R. and Ray,S.C. (1991) *Nucleic Acids Res.*, **19**, 3185–3192.

Chan,E.K.L., Hamel,J.C., Buyon,J.P. and Tan,E.M. (1991) *J. Clin. Invest.*, **87**, 68–76.

Chen,Z., Chen,S.-J., Tong,J.-H., Zhu,Y.-J., Huang,M.-E., Wang,W.-C., Wu,Y., Sun,G.-L., Wang,Z.-Y., Larsen,C.-J. and Berger,R. (1991) *Leukemia*, **4**, 288–292.

Chomienne,C., Ballerini,P., Balitrand,M., Bernard,J.F., Boivin,P., Daniel,M.T., Berger,R., Castaigne,S. and Degos,L. (1989) *The Lancet*, **i**, 746–747.

Clarkson,B. (1991) *Cancer Cells*, **3**, 211–219.

Cleary,M.L. (1991) *Cell*, **66**, 619–622.

Cohen,C. and Parry,D.A.D. (1986) *Trends Biochem. Sci.*, **11**, 245–248.

Davidson,A.J. and Scott,J.E. (1986) *Science*, **249**, 157–160.

Evans,R.M. (1988) *Science*, **240**, 889–895.

Forman,B.M. and Samuels,H.H. (1990) *Mol. Endocrinol.*, **4**, 1293–1301.

Freemont,P.S., Hanson,I.M. and Trowsdale,J. (1991) *Cell*, **64**, 483–484.

Frohman,M.A. and Martin,G.R. (1991) *Technique*, **1**, 165–170.

Fu,X.D. and Maniatis,T. (1990) *Nature*, **343**, 437–441.

Gall,J.G., (1991) *Science*, **252**, 1499–1500.

Garnier,J., Osguthorpe,D.J. and Robson,B.J. (1978) *Mol. Biol.*, **120**, 97–120.

Gascuel,O. and Golmard,J.L. (1988) *Comput. Appl. Biosci.*, **4**, 357–365.

Gaub,M.P., Lutz,Y., Ruberte,E., Petkovich,M., Brand,N. and Chambon,P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 3089–3093.

Gelman,I.H. and Silverstein,S.J. (1987) *Virology*, **61**, 2286–2296.

Giguère,V., Ong,E.S., Segui,P. and Evans,R.M. (1987) *Nature*, **330**, 624–629.

Glass,C.K., Devary,O.V. and Rosenfeld,M.G. (1990) *Cell*, **63**, 729–738.

Green,S. and Chambon,P. (1988) *Trends Genet.*, **4**, 309–314.

Green,S., Walter,P., Kumar,V., Krust,A., Bornert,J.M., Argos,P. and Chambon,P. (1986) *Nature*, **320**, 134–139.

Green,S., Issemann,I. and Scheer,E. (1988) *Nucleic Acids Res.*, **16**, 369–370.

Haupt,Y., Alexander,W.S., Barri,G., Klinken,S.P. and Adams,J.M. (1991) *Cell*, **65**, 753–763.

Huang,M.E., Ye,Y.I., Chen,S.R., Chai,J.R., Lu,J.X., Zhoa,L., Gu,L.J. and Wang,Z.Y. (1988) *Blood*, **72**, 567.

Itoh,K., Itoh,Y. and Frank,M.B. (1991) *J. Clin. Invest.*, **87**, 177–186.

Jones,S., Weber,S. and Prakash,L. (1988) *Nucleic Acids Res.*, **16**, 7119–7131.

Kakizuka,A., Miller,W.H.,Jr, Umesono,K., Warrell,R.P.,Jr, Frankel,S.R., Murty,V.V.V.S., Dimitrovsky,E. and Evans,R.M. (1991) *Cell*, **66**, 663–674.

Kastner,P., Krust,A., Mendelsohn,C., Garnier,J.-M., Zelent,A., Leroy,P., Staub,A. and Chambon,P. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2700–2704.

Kastner,P., Brand,N., Krust,A., Leroy,P., Mendelsohn,C., Petkovich,M., Zelent,A. and Chambon,P. (1991) In Hincliffe,J.R., Hurle,J. and Summerbell,D. (eds), *Developmental Patterning of the Vertebrate Limb*. NATO ASI Series, Plenum Press, London, in press.

Lanotte,M., Martin-Thouvenin,M., Najman,S., Ballerini,P., Valensi,S. and Berger,R. (1991) *Blood*, **77**, 1080–1086.

Leid,M. *et al.* (1992) *Cell*, in press.

Leroy,P., Krust,A., Zelent,A., Mendelsohn,C., Garnier-J.-M., Kastner,P., Dierich,A. and Chambon,P. (1991a) *EMBO J.*, **10**, 59–69.

Leroy,P., Nakshatri,H. and Chambon,P. (1991b) *Proc. Natl. Acad. Sci. USA*, **88**, 10138–10142.

van Lohuizen,M., Verbeek,S., Scheijen,B., Wientjens,E., van der Gulden,H. and Berns,A. (1991a) *Cell*, **65**, 737–752.

van Lohuizen,M., Frasch,M., Wientjens,E. and Berns,A. (1991b) *Nature*, **353**, 353–355.

Loh,E.Y., Elliott,J.F., Cwirla,S., Lanier,L.L. and Davis,M.M. (1989) *Science*, **243**, 217–220.

Longo,L., Pandolfi,P.P., Biondi,A., Rambaldi,A., Mencarelli,A., Lo Coco,F., Diverio,D., Pegoraro,L., Avanzi,G., Tabilio,A., Zangrilli,D., Alcalay,M., Donti,E., Grigani,F. and Pelicci,P.G. (1990) *J. Exp. Med.*, **172**, 1571–1575.

Lübbert,M. and Koeffler,H.P. (1988) *Cancer Rev.*, **10**, 33–62.

Lupas,A., van Dyke,M. and Stock,J. (1991) *Science*, **252**, 1162–1164.

Lutz,Y., Jacob,M. and Fuchs,J.P. (1988) *Exp. Cell Res.*. **175**, 109–124.

Mattei,M.G., Petkovich,M., Mattei,J.F., Brand, N. and Chambon,P. (1988) *Hum. Genet.*, **80**, 186–188.

Mendelsohn,C., Ruberte,E., LeMeuer,M., Morriss-Kay,G. and Chambon,P. (1991) *Development*, **113**, 723–734.

Miki,T., Fleming,T.P., Crescenzi,M., Molloy,C.J., Blam,S.B., Reynolds,S.H. and Aaronson,S.A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 5167–5171.

Mitchell,P.J. and Tjian,R. (1989) *Science*, **245**, 371–378.

Nagpal,S. and Ostrove,J.M. (1991) *J. Virology*, **65**, 5289–5296.

Pandolfi,P.P., Grignani,F., Alcalay,M., Mencarelli,A., Biondi,A., LoCoCo,F., Grignani,F. and Pelicci,P.G. (1991) *Oncogene*, **6**, 1285–1292.

Patarca,R., Schwartz,J., Singh,R.P., Kong,Q-T., Murphy,E., Anderson,Y., Sheng,F.-Y.W., Singh,P., Johnson,K.A., Guarnagia,S.M., Durfee,T., Blattner,F. and Cantor,H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2733–2737.

Petkovich,M., Brand,N.J., Krust,A. and Chambon,P. (1987) *Nature*, **330**, 444–450.

Rochette-Egly,C., Lutz,Y., Saunders,M., Scheuer,I., Gaub,M.P. and Chambon,P. (1991) *J. Cell. Biol.*, **115**, 535–545.

Rowley,J.D., Golomb,H.M., and Dougherty,C. (1977) *The Lancet*, **i**, 549.

Sawyers,C.L., Denny,C.T. and Witte,O.N. (1991) *Cell*, **64**, 337–350.
Schatz,D.G., Oettinger,M.A. and Baltimore,D. (1989) *Cell*, **59**, 1035–1048.
Silver,P.A. (1991) *Cell*, **64**, 489–497.
Smith,W.C., Nakshatri,H., Leroy,P., Rees,J. and Chambon,P. (1991) *EMBO J.*, **10**, 2223–2230.
Tagawa,M., Sakamoto,T., Shigemoto,K., Matsubara,H., Tamura,Y., Ito,T., Nakamura,I., Okitsu,A., Imai,K. and Taniguchi,M. (1990) *J. Biol. Chem.*, **265**, 20021–20026.
Takahashi,M., Inaguma,Y., Hiai,H. and Hirose,F. (1988) *Mol. Cell. Biol.*, **8**, 1853–1856.
Tasset,D. Tora,L., Fromental,C., Scheer,E. and Chambon,P. (1990) *Cell*, **62**, 1177–1187.
de Thé,H., Chomienne,C., Lanotte,M., Degos,L. and Dejean,A. (1990) *Nature*, **347**, 558–561 .
de Thé,H., Lavau,C., Marchio,A., Chomienne,C., Degos,L. and Dejean,A. (1991) *Cell*, **66**, 675–684.
Tora,L., White,J., Brou,C., Tasset,D., Webster,N., Scheer,E. and Chambon,P. (1989) *Cell*, **59**, 477–487.
Vallee,B.L., Coleman,J.E. and Auld,D.S. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 999–1003.
Warrel,R.P.,Jr, Frankel,S.R., Miller,W.H.,Jr, Scheinberg,D.A., Itri,L.M., Hittelman,W.N., Vyas,R., Andreeff,M., Tafuri,A., Jakubowski,A., Gabrilove,J., Gordon,M.S. and Dmitrovsky,E. (1991) *New England J. Med.*, **324**, 1385–1393.
Zelent,A., Krust,A., Petkovich,M., Kastner,P. and Chambon,P. (1989) *Nature*, **339**, 714–717.
Zelent,A., Mendelsohn,C., Kastner,P., Krust,A., Garnier,J.-M., Ruffenach,F., Leroy,P. and Chambon,P. (1991) *EMBO J.*, **10**, 71–81.

Received on September 13, 1991; revised on November 5, 1991

Note added in proof

The Myl cDNA sequence has been deposited in the EMBL Data Library under the accession number X63131.