The EMBO Journal Vol.17 No.22 pp.6723–6729, 1998

Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation

Hui Ge[1], Yuanzheng Si and Robert G.Roeder[2]

Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, Bethesda, MD 20892 and [2]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA

[1]Corresponding author
e-mail: geh@box-g.nih.gov

Transcriptional activation in human cell-free systems containing RNA polymerase II and general initiation factors requires the action of one or more additional coactivators. Here, we report the isolation of cDNAs encoding two novel human transcriptional coactivators (p52 and p75) that are derived from alternatively spliced products of a single gene and share a region of 325 residues, but show distinct coactivator properties. p52 and p75 both show strong interactions with the VP16 activation domain and several components of the general transcriptional machinery. p52, like the previously described PC4, is a potent broad-specificity coactivator, whereas p75 is less active for most activation domains. These results suggest that p52 is a general transcriptional coactivator that mediates functional interactions between upstream sequence-specific activators and the general transcription apparatus, possibly through a novel mechanism.

Keywords: coactivators/p52/p75/PC4/transcriptional activation

Introduction

Modulation of transcriptional activation by RNA polymerase II (RNA pol II) in eukaryotic cells is a complex multistep process controlled by at least three distinct classes of transcription factors. The first class includes the general transcription factors, TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH, in addition to RNA pol II, and mediates accurate transcription initiation through common core promoter elements (reviewed by Orphanides *et al*., 1996; Roeder, 1996). The second class consists of gene-specific regulators that bind to DNA elements distal to core promoter elements and regulate the rate of transcription by the general transcription apparatus (Ptashne and Gann, 1990; Hahn, 1993; Pugh, 1996). The third class is a diverse and more recently identified group of cofactors, including both coactivators and corepressors, that are essential for, or modulate, functional interactions between DNA-bound gene-specific regulators and the general transcription factors. [These include gene-specific cofactors associated with DNA-binding regulatory factors (Luo *et al*., 1992; Chrivia *et al*., 1993), cofactors associated with the basal transcriptional machinery (Bjorklund and Kim, 1996; Verrijzer and Tjian, 1996) and various soluble cofactors (Kaiser and Meisterernst, 1996).]

Within the general transcriptional machinery, the TATA box-binding protein (TBP)-associated factors (TAF_{II}s) associated with TFIID have been regarded as coactivators largely on the basis of studies showing that, in purified systems, TFIID and TBP could both mediate basal transcription on TATA-containing promoters whereas only TFIID could mediate activated transcription (reviewed in Roeder, 1996). Consistent with these results and earlier studies of TFIID–activator interactions (Roeder, 1996), specific activator–TAF interactions were correlated with TFIID binding and activator function (Verrijzer and Tjian, 1996). More recent studies in both yeast and human have indicated that TAF_{II}s may not be generally required for activator function (Moqtaderi *et al*., 1996; Walker *et al*., 1996; Oelgeschlager *et al*., 1998) and that they are involved in core promoter-selective basal transcription functions (Nakatani *et al*., 1990; Martinez *et al*., 1994; Verrijzer *et al*., 1995; Burke and Kadonaga, 1996, 1997; Shen and Green, 1997).

Other coactivators associated with the general transcriptional machinery are the genetically and biochemically defined complex of factors (including SRBs, MEDs and Gal11) that can be isolated, in yeast and human, in association with RNA pol II (Kim *et al*., 1994; Koleske and Young, 1994; Bjorklund and Kim, 1996; Kaiser and Meisterernst, 1996; Orphanides *et al*., 1996). These coactivator complexes may facilitate RNA pol II recruitment through interactions with activators (Koleske and Young, 1995; Ptashne and Gann, 1997) and may mediate activator function in the absence of TAF_{II}s (Kim *et al*., 1994; Koleske and Young, 1994; Oelgeschlager *et al*., 1998).

An apparently distinct set of general coactivators, termed positive cofactors (PCs) has been identified in human HeLa cells. At least four PCs (PC1, PC2, PC3 and PC4) have been separated and completely or partially purified from the upstream stimulatory activity (USA) fraction, while two less well characterized PCs (PC5 and PC6) have been found in other HeLa cell nuclear extract-derived fractions (Kaiser and Meisterernst, 1996). Although the various PCs can function independently, synergistic functions in enhancing the absolute levels of transcription and/or the fold-induction in response to an activator have been reported (Halle *et al*., 1995; Geurmah *et al*., 1998; Luo *et al*., 1998). The best characterized PC, PC4, is a single- and double-stranded DNA-binding protein that mediates activator-dependent transcription in a TBP- and TAF-dependent manner, but is not required for basal activity in an *in vitro* reconstituted transcription system. It acts as a general transcriptional coactivator for a variety of activators and, consistent with its role as an adaptor, directly interacts both with activation domains of regu-

latory factors and with the general transcription factor TFIIA. All these activities of PC4 are regulated negatively *in vivo* by phosphorylation (Ge and Roeder, 1994a; Ge *et al*., 1994; Kretzschmar *et al*., 1994a). Recent studies have indicated that PC4 may also interact with RNA pol II, and have further suggested a dynamic model for PC4 coactivator action that involves an inactive intermediate in PIC assembly and function (Malik *et al*., 1998).

The overall complexity of coactivators, including the lack of an absolute requirement for specific coactivators in various *in vivo* and *in vitro* assays, raises questions of possible functional redundancies and activator specificities not only for the TAF_{II}s and SRBs but also for the PCs. This report describes the cloning and characterization of two closely related novel coactivators (p52 and p75), present in USA-derived PC4-containing fractions, that mediate activator-dependent, but not basal, transcription by various activators. In agreement with an adaptor model for their coactivator functions, p52 and p75 both interact directly with the VP16 activation domain and with components of the general transcription machinery.

Results

Isolation of cDNAs encoding novel proteins p52 and p75

During the course of purifying the general transcriptional coactivator PC4, an ~75 kDa polypeptide (p75) co-purified with PC4 activity (Ge and Roeder, 1994a). Microsequence analyses of the excised polypeptide revealed that the N-terminal region of p75 showed a high degree of similarity to human hepatoma-derived growth factor (HDGF), a protein that was identified originally in a human hepatoma-derived cell line but is ubiquitously expressed in several other cell lines and tissues (Nakamura *et al*., 1994). A degenerate oligonucleotide based on the N-terminal sequence was used to isolate, from a HeLa cell cDNA library, a clone containing a 1.8 kb insertion. The cDNA contained an open reading frame (ORF; nucleotides 78–1076) capable of encoding a protein of 333 amino acid residues (Figure 1A). The coding region contained all of the determined peptide sequences, including one from the N-terminus and two from internal trypsin-digested fragments (underlined in Figure 1A), and there were two polyadenylation [poly(A)] signals (AAUAAA) downstream of the stop codon. However, the predicted molecular mass (37.5 kDa) is smaller than the size of the authentic protein estimated from SDS–gel analysis (75 kDa), and expression of the cDNA either in *Escherichia coli* (see Figure 3A) or by *in vitro* translation (data not shown) gave rise to a 52 kDa (p52) product as estimated by SDS–PAGE. Thus, it appeared that the isolated p52 cDNA might be derived from an alternatively spliced product of the p75 precursor RNA. Indeed, a rescreen of the HeLa cell cDNA library with a 3′ cDNA probe (see Materials and methods) yielded a group of 10 positive clones with sequence divergence from the p52 cDNA starting at a position 22 bp upstream of the stop codon, thereby generating an extended ORF of 530 amino acid residues (Figure 1B). A bacterially expressed protein (p75) encoded by one of these cDNAs (clone 7) exhibited a mobility in SDS–PAGE similar to that of the native 75 kDa protein (see Figure 3A).

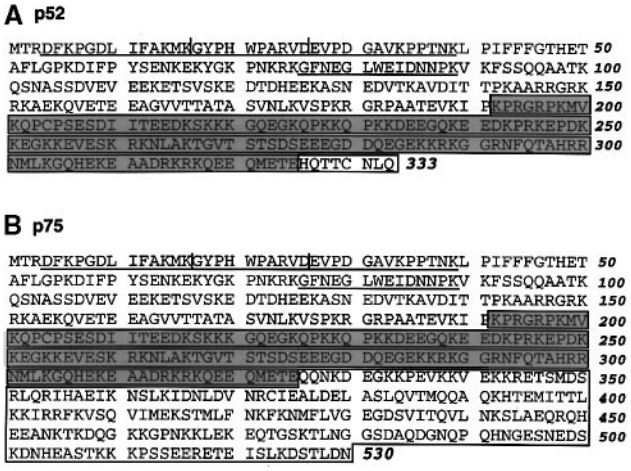

Fig. 1. Sequence analyses of p52 and p75. (**A**) Deduced amino acid sequence of p52. The amino acid sequences obtained from microsequencing of N-terminal (residues 4–26) and internal (residues 17–39 and 76–89) peptides are underlined. The highly charged C-terminal domain is shaded. (**B**) Deduced amino acid sequence of p75. The boxed amino acid residues indicate the C-terminal region unique to p75. These protein sequences have been submitted to the DDBJ/EMBL/GenBank under the accession numbers AF098482 and AF098483.

p52 and p75 are both highly charged proteins, with lysine, arginine, glutamate and aspartate comprising 40.5 and 39.2%, respectively, of the total residues. Most notable is a common 134 residue region, at the C-terminus of p52 and internal to p75, that contains 71 (53%) charged residues. In addition to the striking similarity between the N-terminus of p52 and p75 and that of the human HDGF protein (71% identity within 100 amino acid residues; data not shown), the C-terminal region of p52 (also present in p75) shows similarity to human HMG-1, a multifunctional non-histone chromatin protein involved in many steps of gene regulation (Bustin and Reeves, 1996; see Discussion).

Northern blot analyses of mRNAs from human tissues revealed three major bands of 3.4, 2.8 and 1.8 kb with a probe corresponding to a C-terminal fragment of the p52-coding region (Figure 2A) and two major bands of 3.4 and 2.8 kb with a probe corresponding to a C-terminal fragment of the p75-coding region (Figure 2B); this indicates that the smallest species of mRNA of 1.8 kb corresponds to p52, whereas the other two larger species of mRNA of 3.4 and 2.8 kb correspond to p75. p52 and p75 mRNAs are both ubiquitously expressed but the p52 mRNA is most abundant in the testis (Figure 2A, lane 12), followed by thymus and brain (lanes 2 and 10), whereas the p75 mRNA is most abundant in thymus (Figure 2B, lane 10).

Functional analysis of recombinant p52 and p75

Since p75 co-purified with the transcriptional coactivator PC4, the potential roles of p52 and p75 in transcriptional regulation were examined. Both cDNAs were introduced into a pET vector that introduced a six histidine tag and a heart muscle kinase (HMK) site at the N-terminus of each. Bacterially expressed proteins were purified (Figure 3A) by sequential chromatography on nickel affinity, Mono S and Superdex 200 columns. Polyclonal antibodies against recombinant p52 recognized corresponding recom-

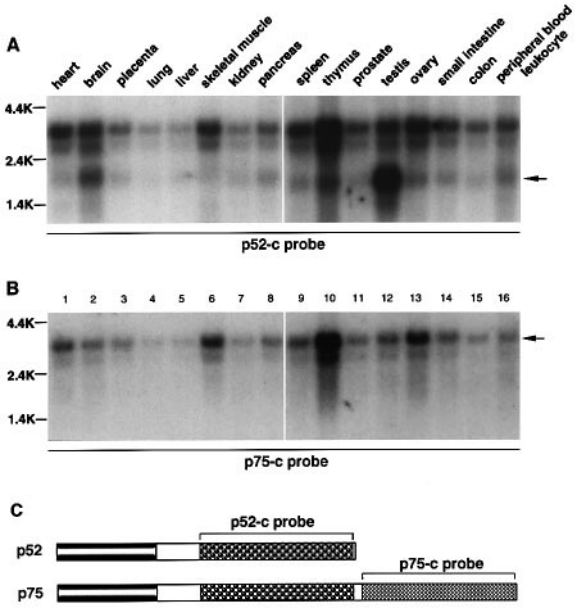

Fig. 2. Northern blot analyses. (**A**) Poly(A)$^+$ RNAs isolated from human tissues (ClonTech) were hybridized with the 3′ p52 cDNA probe (*Pst*I–*Bgl*II fragment, bp 624–1160). The sources of RNAs are indicated on the top, and the sizes of RNA markers are indicated (in kb) on the left. (**B**) The filters from (A) were stripped and reprobed with a PCR fragment corresponding to the unique 3′ coding region of p75. (**C**) Schematic representation of p52 and p75 protein structures and probes used for Northern blot analyses.

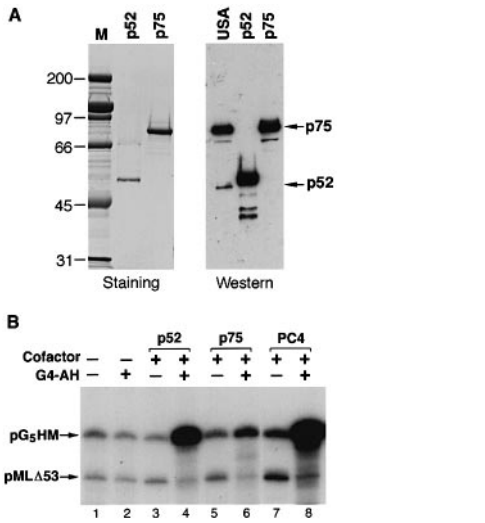

Fig. 3. Functional analyses of recombinant p52 and p75. (**A**) Recombinant p52 [6H(K)p52] and p75 [6H(K)p75] were analyzed by SDS–PAGE and visualized by either Coomassie Blue staining (left panel) or immunoblot using polyclonal antibodies generated from full-length recombinant p52 protein (right panel). Authentic p52 and p75 present in the partially purified PC4-containing fraction (USA) were also detected by anti-p52 antibodies and have mobilities similar to those of the recombinant proteins. Protein markers (M) are indicated in kDa on the left. (**B**) Functional analyses. Coactivator activities of recombinant p52, p75 and PC4 were tested in an *in vitro* transcription assay reconstituted with purified factors either in the presence (+) or in the absence (–) of activator GAL4-AH as indicated. Transcripts of pG5HMC2AT (activated template) and pMLΔ53 (control basal template) are indicated by arrows.

binant p52 and p75 proteins, as well as natural p52 and p75 proteins in the partially purified PC4-containing USA coactivator fraction (Figure 3A, right panel). Thus the two cloned cDNAs indeed encode the native p52 and p75 proteins.

The coactivator functions of p52 and p75 were assessed in a standard *in vitro* transcription system reconstituted with highly purified general transcription factors (Ge *et al*., 1996). Previous studies showed that this system requires additional cofactors, either USA or derived components (Meisterernst *et al*., 1991), and that PC4 alone can markedly enhance transcription by diverse activators (Ge and Roeder, 1994a; Kretzschmar *et al*., 1994a). PC4 marginally stimulated basal level transcription in the absence of activator GAL4-AH (Figure 3B, lane 7), but markedly enhanced activated transcription on the pG5HMC2AT reporter template (containing five GAL4 sites) in the presence of GAL4-AH (Figure 3B, lane 8). Like recombinant PC4, recombinant p52 and p75 had little or no effect on transcription in the absence of GAL-AH (Figure 3B, lanes 3 and 5). However, like PC4, p52 greatly enhanced transcription in the presence of GAL4-AH (Figure 3B, lane 4). p75 also enhanced transcription in the presence of GAL4-AH (Figure 3B, lane 6), but its effect was minimal (~3-fold) compared with p52 or PC4 (>15-fold). These results indicate that recombinant p52 (and to a lesser extent p75) is a transcriptional coactivator capable of substituting for PC4 to potentiate GAL4-AH-dependent transcription *in vitro*.

p52 and p75 serve as coactivators for different activators

Given that recombinant p52 and p75 (particularly p52) both can facilitate transcriptional activation by GAL4-AH, which contains an artificial acidic activation domain (Giniger and Ptashne, 1987; Wang *et al*., 1992), we asked whether, like PC4, they also could function as more general coactivators to potentiate activated transcription by other activators. Indeed, as shown in Figure 4A, both p52 and p75 significantly enhanced activation both by the acidic activation domain of VP16 (lanes 2–7) and by the acidic activation domain (Martin *et al*., 1990) of the pseudorabies immediate early protein (IE) (lanes 12–17) in a concentration-dependent manner. The further analysis of Figure 4B shows that recombinant p52 strongly stimulates transcriptional activation by GAL4 fusion proteins containing the proline-rich activation domain of CTF (Figure 4B, lane 6 versus lane 2), the glutamine-rich activation domain of Sp1 (lane 7 versus lane 3), the activation domain of adenovirus E1A (lane 9 versus lane 5) and, as shown above, the IE activation domain (lane 8 versus lane 4).

The coactivator functions observed with p52 closely parallel those of PC4 in the same assay (Figure 4B, lanes 6–9 versus lanes 14–17; see Figure 4C for quantification). In contrast, p75 has only a moderate effect on activation by the proline-rich activation domain of CTF and (as indicated above) the acidic activation domain of pseudorabies IE (Figure 4B, lanes 10 and 12) and does not significantly enhance activation by either the glutamine-rich activation domain of Sp1 or the activation domain of E1A (Figure 4B, lanes 11 and 13). Similar results have been obtained when purified authentic Sp1 protein is tested in this system (Ge *et al.*, 1998). Taken together, these observations demonstrate that recombinant p52 protein can act as a general transcriptional coactivator, comparable

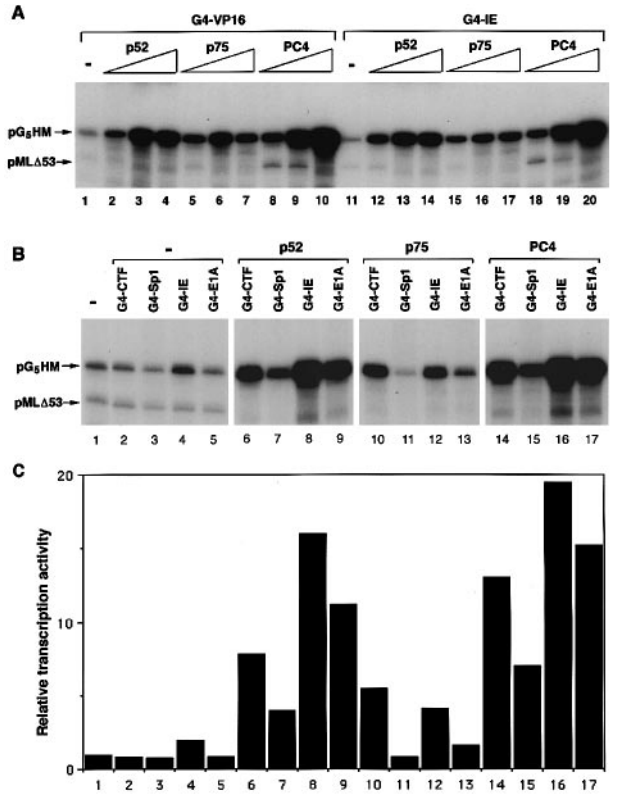

Fig. 4. Activator specificities of p52 and p75. (**A**) *In vitro* transcription assays were performed with increasing amounts (1.5, 3 and 4.5 pmol) of recombinant p52, p75 or PC4 in the presence of either GAL4-VP16 (lanes 1–10) or GAL4-IE (lanes 11–20). (**B**) Additional GAL4-fused activation domains, including GAL4-CTF (lanes 2, 6, 10 and 14), GAL4-Sp1 (lanes 3, 7, 11 and 15), GAL4-E1A (lanes 5, 9, 13 and 17) as well as GAL4-IE (lanes 4, 8, 12 and 16), were tested for the specificities of p52, p75 and PC4. (**C**) Quantitative representation of coactivator activities by different types of activation domains. Only the transcript from the activated template (pG5HMC2AT) in each reaction was determined. The relative transcription activity from lane 1 (absence of activator and coactivator) was normalized as 1.

with PC4, whereas p75 functions less actively in potentiating activator function.

p52 and p75 directly bind to the VP16 activation domain

As transcriptional coactivators, p52 and p75 could serve as adaptors for functional interactions between activation domains of sequence-specific activators and component(s) of the basal transcription machinery. To investigate this possibility, and based on previous studies of PC4 (Ge and Roeder, 1994a), we assessed the binding of p52 or p75 to immobilized GST–VP16 fusion proteins that contained the fully active bipartite activation domain encompassing VP16 residues 413–490, a partially active domain lacking the C-terminal 34 residues (Δ456) or an inactive domain containing an additional phenyalanine to proline point mutation at position 442 in the truncated derivative (Δ456FP442) (Triezenberg *et al*., 1988; Cress and Triezenberg, 1990; Lin and Green, 1991). Recombinant GST fusion proteins (Figure 5A) were expressed in *E.coli* and purified by glutathione–Sepharose. After incubation of ^{32}P-labeled p52 or p75 with GST–VP16 fusion proteins and washing with BC buffer containing 200 mM KCl, the bound proteins were analyzed by SDS–PAGE. As shown

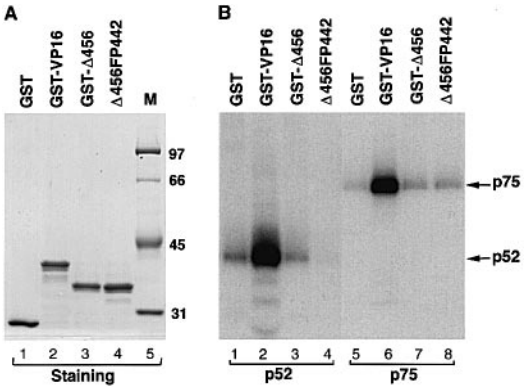

Fig. 5. p52 and p75 directly interact with the VP16 activation domain. (**A**) Purified GST fusion proteins, GST–VP16 (wild-type), GST–Δ456 (C-terminal deletion of VP16 activation domain) and Δ456FP442 (point mutation of the truncated version), were analyzed by SDS–PAGE and visualized by Coomassie Blue staining. (**B**) Protein–protein interaction assays for direct binding of p52 and p75. ^{32}P-labeled recombinant p52 (lanes 1–4) or p75 (lanes 5–8) was incubated with an immobilized GST–VP16 activation domain fusion protein as indicated. The bound proteins were analyzed by SDS–PAGE and autoradiography. Approximately 20–25% of the input p52 or p75 was retained on the GST–VP16 resin.

in Figure 5B, p52 and p75 both bound strongly to GST–VP16 (lanes 2 and 6) but only very weakly, and at levels close to the background levels observed with GST (lanes 1 and 5), to GST–Δ456 (lanes 3 and 7) and GST–Δ456FP442 (lanes 4 and 8). Thus the function of the VP16 activation domain in a p52/p75-dependent assay correlates well with its ability to bind p52/p75.

Previous studies indicated that the potent coactivator function of human PC4 may be mediated by direct interactions both with various activators and with TFIIA (Ge and Roeder, 1994a; Kretzschmar *et al*., 1994a), and potentially RNA pol II (Malik *et al*., 1998), whereas the putative yeast homolog of PC4 showed interactions with TFIIB (Henry *et al*., 1996; Knaus *et al*., 1996). To test possible interactions between p52 or p75 and components of the basal transcription machinery, we analyzed interactions of recombinant GST–p52 or GST–p75 with natural proteins in HeLa nuclear extract. Unexpectedly, all tested general transcription factors were bound to the GST–p52 (but only few to GST–p75) column at the levels significantly above the background levels observed for GST alone (data not shown). It seems unlikely that all these interactions are functionally relevant. Although in a separate experiment we have found that interactions between p52/p75 and some of GST-fused general transcription factors (RAP74 and two subunits of RNA pol II) are relatively stronger than others, we cannot conclude at present that these interactions specifically mediate p52/p75 coactivator function. Therefore, further characterization is still required to address a fundamental interaction(s) between p52/p75 and the general transcriptional machinery.

Discussion

Our previous studies using an *in vitro* system reconstituted with purified general transcription factors have demonstrated a requirement for PC4 (or other positive cofactors), in addition to TAF_{II}s, for maximal levels of activator-

dependent transcription of class II genes. The present report describes two novel coactivators (p52 and p75), generated from alternatively spliced products of a single gene, that mediate activator-dependent transcription by diverse activators. Although p75 contains all but the C-terminal eight residues of p52, along with a unique region of 205 residues, it is less efficient than p52 in mediating activated transcription. Like PC4, p52 and p75 appear to serve as adaptors between sequence-specific activators and the basal transcription machinery, but possibly through the interactions with different general factors.

Relationship of p52 and p75 to the general transcriptional coactivator PC4

p52 and p75 resemble PC4 with respect to their ability to function as general coactivators (mainly p52) and their ability to interact both with a functional activation domain and with components of the general transcriptional machinery. Hence, like PC4, they may both serve as adaptors between these two classes of factors, but other considerations suggest that the details of their mechanisms of action may be different. First, they differ completely in sequence from PC4 (Ge and Roeder, 1994a; Kretzschmar *et al*., 1994a). Secondly, their primary targets in the general transcriptional machinery may be different. Thus, whereas PC4 has been shown to interact directly with TFIIA and TFIIA–TBP–promoter complexes (Ge and Roeder, 1994a; Kretzschmar *et al*., 1994a; Kaiser *et al*., 1995), and possibly with RNA pol II (Malik *et al*., 1998), p52 and p75 show interactions with several factors in the general transcriptional machinery and the strongest interactions with TFIIF and RNA pol II subunits in a GST pull-down assay (data not shown). The latter point is particularly intriguing in view of the observations that TFIIF stably associates with RNA pol II and that these components function not only in PIC assembly, but also during the initiation, elongation and reinitiation stages of transcription (Flores *et al*., 1991; Tan *et al*., 1995; Forget *et al*., 1997). Thus, p52 (and/or p75) functions could be more important in both early and later events in the transcription cycle, whereas PC4 may act in early stages to facilitate TFIID recruitment and initiation (Halle *et al*., 1995; Malik *et al*., 1998). Interestingly, we have also found that both p52 and p75 preferentially interact with the dephosphorylated (active) form of PC4 (unpublished data). Major unanswered questions concern the possibility of cooperativity between p52/p75 and PC4, as well as structural features important for activator versus general factor interactions and for the activator selectivity of p52 versus p75.

Relationship of p52 and p75 with other positive cofactors

Although the TAF_{II} components of TFIID are proving to have increasingly complex and variably required functions as coactivators and core promoter selectivity factors (Introduction), past and present studies have shown that they are essential coactivators in cell-free systems reconstituted with highly purified factors and, conversely, that positive cofactors such as PC4 (Ge and Roeder, 1994a; Kretzschmar *et al*., 1994a) and p52/p75 (Figures 3B and 4) are still essential in these systems even when TAF_{II}s are present. Hence, it is clear that PC4 and p52/p75 function differently from TAF_{II}s, but the exact nature of the interdependent functions remains to be established.

Apart from PC4, additional positive cofactors (PC1, PC2, PC3/DR2/Topo l, PC5 and PC6) have been shown to enhance activator function in conjunction with TAF_{II}s and other components of the general transcriptional machinery (reviewed in Kaiser and Meisterernst, 1996). Although p52 and p75 are derived from the same USA fraction as positive cofactors 1–4, they clearly differ from PC1 (Meisterernst *et al*., 1997), PC3 (Kretzschmar *et al*., 1993; Merino *et al*., 1993) and PC4 (above) in primary sequence. In addition, they also appear to differ from PC2 (Kretzschmar *et al*., 1994b), PC5 and PC6 (Kaiser and Meisterernst, 1996) on the basis of chromatographic and functional properties. However, until these latter factors, especially the ~500 kDa PC2 complex, are purified and cognate cDNAs cloned, potential relationships of p52 and p75, possibly as subcomponents, cannot be ruled out.

Non-histone chromosomal proteins HMG-1 and -2 also have been implicated in regulation of transcriptional activation. HMG-1 can affect transcription either positively or negatively by direct interactions with sequence-specific DNA-binding proteins (Onate *et al*., 1994; Ge and Roeder, 1994b; Zappavigna *et al*., 1996; Jayaraman *et al*., 1998), whereas HMG-2 protein can also act both as a negative cofactor to repress basal transcription (Stelzer *et al*., 1994) and as a positive cofactor in response to sequence-specific activators in defined *in vitro* systems (Shykind *et al*., 1995; Zwilling *et al*., 1995). Interestingly, like HMG-1 and HMG-2, p52 and p75 each contain (in their common region) an N-terminal HMG box-like structure (~20% identity and 50% positive within N-terminal residues 6–185, data not shown) and a highly charged C-terminal domain (data not shown). However, it is still unclear whether p52 and p75 mediate transcriptional activation in a similar fashion to that of HMG proteins (cf. Shykind *et al*., 1995).

Significance of p52 and p75 interactions with multiple target proteins

A number of activators and coactivators, as exemplified by VP16 (reviewed by Greenblatt and Ingles, 1996), GAL4 (Ptashne and Gann, 1990), p53 (Levine, 1997) and PC4 (Ge and Roeder 1994a; Malik *et al*., 1998), have been shown to interact with multiple components of the general transcriptional machinery, and in several cases the functional significance of these interactions has been indicated by mutational analyses. Reciprocally, many general transcriptional factors, as exemplified by TBP and TFIIB (reviewed in Choy and Green, 1993; Burley and Roeder, 1996), have been shown to interact with multiple activators. The presence of multiple targets in the general transcription machinery for the same or different activators/coactivators may serve as the basis for cooperative functions of these components (Ptashne and Gann, 1997; Carey, 1998), whereas multiprotein interactions, in general, may provide more efficient, multistep regulatory paradigms for gene activation.

In this regard, it may be significant that p52 and p75 show interactions not only with activation domains of sequence-specific activators, but also with several components of the general transcriptional machinery (data not shown) although it is unclear which interaction is specific

and functionally relevant. A potential explanation for these observations is that the highly charged C-terminal domain of p52 (71 basic and acidic amino acids in the C-terminal 134 residues) contains a number of distinct sites for the various (multi)protein interactions, thereby contributing distinct functions for gene activation processes. Such a multifunctional structure has been proposed for the C-terminal domain (CTD) of the largest subunit of RNA pol II, which contains 17–52 (depending on the organism) heptapeptide repeats. Apart from important roles in transcriptional regulation through interactions with general transcription factors (Dahmus, 1996) and the SRB-containing mediator complex of the RNA pol II holoenzyme (Koleske and Young, 1995; Bjorklund and Kim, 1996), the CTD also plays a role in linking transcription to pre-mRNA processing [including 3′ cleavage/poly(A) and pre-mRNA splicing] (reviewed by Steinmetz, 1997; Corden and Patturajan, 1997 and references therein).

Materials and methods

Isolation of cDNAs encoding p52 and p75

p75 was co-purified with the general transcriptional coactivator PC4 as previously described (Ge and Roeder, 1994a). The p75-containing Mono S fraction was resolved by SDS–PAGE and transferred to a nitrocellulose membrane. After visualization by Ponceau S staining, the 75 kDa polypeptide was excised and subjected to N-terminal sequencing and *in situ* trypsin digestion for internal sequence analyses (by the Protein/DNA Technology Center at the Rockefeller University). One N-terminal (XXDFKPG-DLIFAKMKGYPHXPAXVD) and two internal (G/KYPT/HSPAS/RVDEVPDG/AAVKPPTNK and GFNEGLWEIDNNPK) sequences were obtained. A degenerate oligonucleotide, 5′-GATTTCAARCCIGGIGAT-CTITTTGCIAARATGAARGGITACCCICA-3′, based on the N-terminal peptide sequence according to human codon bias, was used to screen a HeLa cDNA library in the λZAPII vector. One of the resulting three positive clones contained a 1.8 kb insertion with a 333 amino acid ORF (Figure 1A) that represents an alternatively spliced isoform of p75 designated p52. A second screen of the HeLa cDNA library with the 3′ coding region of p52 (*Pst*I–*Bgl*II fragment, from the coding sequence for amino acid residue 184 to 85 bp downstream of the stop codon) yielded 10 positive clones whose inserts have 3′ untranslated region (UTR) sequences distinct from that of p52 cDNA. Although most have a long 3′ UTR and poly(A) tails, none contain 5′ coding regions. The 5′ end of a 3.4 kb insertion (clone 7) corresponds to the sequence of p52 cDNA starting at bp 620, 5 bp upstream of a unique *Pst*I site, but the sequence diverges from bp 1054, 22 bp upstream of the p52 stop codon. This different 3′ sequence generates an extended 530 amino acid ORF which was demonstrated to be the cDNA encoding p75 (Figure 1B). The 5′ region of p75 was confirmed by PCR using a 5′ primer corresponding to 5′ UTR and start codon sequences of p52 cDNA and a 3′ primer corresponding to a unique sequence of p75 cDNA. Northern blot analyses of poly(A)$^+$ RNAs isolated from human tissues (ClonTech) were performed according the manufacturer's instructions using the 3′ half of the p52-coding region (*Pst*I–*Bgl*II fragment) as the p52-C probe (Figure 2A), and a 610 bp PCR fragment corresponding to the 3′ coding region of p75 as the p75-C probe (Figure 2B).

Expression of recombinant p52 and p75

The *Nru*I–*Eco*RV fragment of p52 cDNA was introduced into the *Sma*I site of pGEX2T(K) vector (Ge and Roeder, 1994a) to generate plasmid GST–K-p52. The 6H(K)p52 plasmid was generated by inserting an *Eco*RI fragment from GST–K-p52 into the *Eco*RI site of the pET11a-6H(K) vector, which includes sequences encoding six histidines and an HMK site. GST–K-p75 and 6H(K)p75 plasmids were created by replacing the *Pst*I–*Eco*RI fragment (627 bp) from either GST–K-p52 or 6H(K)p52 plasmid with the *Pst*I–*Eco*RI fragment (1675 bp) from p75 cDNA (clone 7).

All these plasmids were transformed into *E.coli* cells and expressed by isopropyl-β-D-thiogalactopyranoside (IPTG) induction as previously described (Ge and Roeder, 1994a). 6H(K)p52 and 6H(K)p75 proteins were purified by Ni^{2+}-NTA–agarose affinity column, and further purified by FPLC Mono S and Superdex 200 chromatography. The GST fusion proteins were purified by a glutathione–Sepharose affinity column.

In vitro transcription assays

In vitro transcription reactions were reconstituted with purified native or recombinant general transcription factors in the presence or absence of GAL4 DNA-binding domain–activation domain fusion proteins as previously described (Ge *et al*., 1996). The activated template (PG5HMC2AT) contains five GAL4 DNA-binding sites upstream of the human immunodeficiency virus type 1 TATA box and adenovirus major late initiator elements linked to a 380 bp G-less cassette. The basal template (pMLΔ53) contains the adenovirus major late core promoter region (–53 to +10) linked to a 300 bp G-less cassette. The relative transcription activity was determined by densitometer (Molecular Dynamics) after ^{32}P-labeled transcripts were analyzed by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography.

Protein–protein interactions

GST pull-down assays were carried out essentially as described (Ge and Roeder, 1994a). In brief, 10–20 μg of each immobilized GST fusion protein was incubated with 10 ng of ^{32}P-labeled (by HMK) 6H(K)p52 or 6H(K)p75 at 4°C for 1 h in buffer A100 containing 20 mM HEPES-Na (pH 7.9), 10% glycerol, 0.2 mM EDTA, 100 mM KCl, 0.5 mM phenylmethylsulfonyl fluoride and 0.1% NP-40 plus 0.5 mg/ml bovine serum albumin. After washing with buffer A200 (200 mM KCl), 20% of the bound proteins were resolved by SDS–PAGE and detected by autoradiography.

Acknowledgements

We thank Drs P.Wade and A.Wolffe for critical comments on the manuscript, Dr Z.Wang for constructs expressing GST-fused subunits of RNA pol II, and X.Zhang for nuclear extract preparation.

References

Bjorklund,S. and Kim,Y.-J. (1996) Mediator of transcriptional regulation. *Trends Biochem. Sci.*, **21**, 335–337.

Burke,T.W. and Kadonaga,J.T. (1996) *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.*, **10**, 711–724.

Burke,T.W. and Kadonaga,J.T. (1997) The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by $TAF_{II}60$ of *Drosophila*. *Genes Dev.*, **11**, 3020–3031.

Burley,S.K. and Roeder,R.G. (1996) Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.*, **65**, 769–799.

Bustin,M. and Reeves,R. (1996) High mobility group proteins. *Prog. Nucleic Acids Res. Mol. Biol.*, **54**, 35–100.

Carey,M. (1998) The enhanceosome and transcriptional synergy. *Cell*, **9**, 5–8.

Choy,B. and Green,M.R. (1993) Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature*, **366**, 531–536.

Chrivia,J.C., Kwok,R.P., Lamb,N., Hagiwara,M., Montminy,M.R. and Goodman,R.H. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature*, **365**, 855–859.

Corden,J.L. and Patturajan,M. (1997) A CTD function linking transcription to splicing. *Trends Biochem. Sci.*, **22**, 413–416.

Cress,W.D. and Triezenberg,S.J. (1990) Critical structural elements of the VP16 transcriptional activation domains. *Science*, **251**, 87–90.

Dahmus,M.E. (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.*, **271**, 19009–19012.

Flores,O., Lu,H., Killeen,M., Greenblatt,J., Burton,Z.F. and Reinberg,D. (1991) The small subunit of transcription factor TFIIF recruits RNA polymerase II into the preinitiation complex. *Proc. Natl Acad. Sci. USA*, **88**, 9999–10003.

Forget,D., Robert,F., Grondin,G., Burton,Z.F., Greenblatt,J. and Coulombe,B. (1997) RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. *Proc. Natl Acad. Sci. USA*, **94**, 7150–7155.

Ge,H. and Roeder,R.G. (1994a) Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. *Cell*, **78**, 513–523.

Ge,H. and Roeder,R.G. (1994b) The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein. *J. Biol. Chem.*, **269**, 17136–17140.

Ge,H., Zhao,Y.M., Chait,B.T. and Roeder,R.G. (1994) Phosphorylation negatively regulates the function of coactivator PC4. *Proc. Natl Acad. Sci. USA*, **91**, 12691–12695.

Ge,H., Martinez,E., Chiang,C.-M. and Roeder,R.G. (1996) Activator-dependent transcription by mammalian RNA polymerase II: *in vitro* reconstitution with general transcription factors and cofactors. *Methods Enzymol*., **274**, 57–71.

Ge,H., Si,Y.Z. and Wolffe,A.P. (1998) A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. *Mol. Cell*, in press.

Giniger,E. and Ptashne,M. (1987) Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. *Nature*, **330**, 670–672.

Greenblatt,J. and Ingles,C.J. (1996) Interaction between acidic transcriptional activation domains of herpes simplex virus activator protein VP16 and transcriptional initiation factor IID. *Methods Enzymol*., **274**, 120–133.

Guermah,M., Malik,S. and Roeder,R.G. (1998) Involvement of TFIID and USA components in transcriptional activation of the human immunodeficiency virus promoter by NF-κB and Sp1. *Mol. Cell. Biol*., **18**, 3234–3244.

Hahn,S. (1993) Efficiency in activation. *Nature*, **363**, 672–673.

Halle,J.P., Stelzer,G., Goppelt,A. and Meisterernst,M. (1995) Activation of transcription by recombinant upstream stimulatory factor 1 is mediated by a novel positive cofactor. *J. Biol. Chem*., **270**, 21307–21311.

Henry,N.L., Bushnell,D.A. and Kornberg,R.D. (1996) A yeast transcriptional stimulatory protein similar to human PC4. *J. Biol. Chem*., **271**, 21842–21847.

Jayaraman,L., Moorthy,N.C., Murthy,K.G.K., Manley,J.L. and Prives,C. (1998) High mobility group protein-1 (HMG-1) is a unique activator of p53. *Genes Dev.*, **12**, 462–472.

Kaiser,K. and Meisterernst,M. (1996) The human general co-factors. *Trends Biochem. Sci*., **21**, 342–345.

Kaiser,K., Stelzer,G. and Meisterernst,M. (1995) The coactivator p15 (PC4) initiates transcriptional activation during TFIIA–TFIID–promoter complex formation. *EMBO J*., **14**, 3520–3527.

Kim,Y.-J., Bjorklund,S., Li,Y., Sayre,M.H. and Kornberg,R.D. (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell*, **77**, 599–608.

Knaus,R., Pollock,R. and Guarente,L. (1996) Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. *EMBO J*., **14**, 1933–1944.

Koleske,A.J. and Young,R.A. (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature*, **368**, 466–469.

Koleske,A.J. and Young,R.A. (1995) The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci*., **20**, 113–116.

Kretzschmar,M., Meisterernst,M. and Roeder,R.G. (1993) Identification of human DNA topoisomerase I as a cofactor for activator-dependent transcription by RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **90**, 11508–11512.

Kretzschmar,M., Kaiser,K., Lottspeich,F. and Meisterenst,M. (1994a) A novel mediator of class II gene transcription with homology to viral immediate-early transcriptional regulators. *Cell*, **78**, 525–534.

Kretzschmar,M., Stelzer,G., Roeder,R.G. and Meisterernst,M. (1994b) RNA polymerase II cofactor PC2 facilitates activation of transcription by GAL4-AH *in vitro*. *Mol. Cell. Biol*., **14**, 3927–3937.

Levine,A.J. (1997) p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323–331.

Lin,Y.-S. and Green,M.R. (1991) Mechanism of action of an acidic transcriptional activator *in vitro*. *Cell*, **64**, 971–981.

Luo,Y., Fujii,H., Gerster,T. and Roeder,R.G. (1992) A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell*, **71**, 231–241.

Luo,Y., Ge,H., Stevens,S., Xiao,H. and Roeder,R.G. (1998) Coactivation by OCA-B: definition of critical regions and synergism with general cofactors. *Mol. Cell. Biol*., **18**, 3803–3810.

Malik,S., Guermah,M. and Roeder,R.G. (1998) A dynamic model for PC4 coactivator function in RNA polymerase II transcription. *Proc. Natl Acad. Sci. USA*, **95**, 2192–2197.

Martin,K.J., Lillie,J.W. and Green,M.R. (1990) Transcriptional activation by the pseudorabies virus immediate early protein. *Genes Dev.*, **4**, 2371–2382.

Martinez,E., Chiang,C.-M., Ge,H. and Roeder,R.G. (1994) TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. *EMBO J*., **13**, 3115–3126.

Meisterernst,M., Roy,A.L., Lieu,H.M. and Roeder,R.G. (1991) Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. *Cell*, **66**, 981–993.

Meisterernst,M., Stelzer,G. and Roeder,R.G, (1997) Poly(ADP-ribose) polymerase enhances activator-dependent transcription *in vitro*. *Proc. Natl Acad. Sci. USA*, **94**, 2261–2265.

Merino,A., Madden,K.R., Lane,W.S., Champoux,J.J. and Reinberg,D. (1993) DNA topoisomerase I is involved in both repression and activation of transcription. *Nature*, **365**, 227–232.

Moqtaderi,Z., Bai,Y., Poon,D., Weil,P.A. and Struhl,K. (1996) TBP-associated factors are not genetically required for transcriptional activation in yeast. *Nature*, **383**, 188–191.

Nakamura,H., Izumoto,Y., Kambe,H., Kuroda,T., Mori,T., Kawamura,K., Yamamoto,H. and Kishimoto,T. (1994) Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. *J. Biol. Chem*., **269**, 25143–25149.

Nakatani,Y., Horikoshi,M., Brenner,M., Yamamoto,T., Besnard,F., Roeder,R.G. and Freese,E. (1990) A downstream initiation element required for efficient TATA box binding and *in vitro* function of TFIID. *Nature*, **348**, 86–88.

Oelgeschlager,T., Tao,Y., Kang,Y.K. and Roeder,R.G. (1998) Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking $TAF_{II}s$. *Mol. Cell*, **1**, 925–931.

Onate,S.A., Prendergast,P., Wagner,J.P., Nissen,M., Reeves,R., Pettijohn,D.E. and Edwards,D.P. (1994) The DNA-bending protein HMG-1 enhances progesterone receptor binding to its target DNA sequences. *Mol. Cell. Biol*., **14**, 3376–3391.

Orphanides,G., Lagrange,T. and Reinberg,D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.*, **10**, 2657–2683.

Ptashne,M. and Gann,A.A.F. (1990) Activators and targets. *Nature*, **346**, 329–331.

Ptashne,M. and Gann,A. (1997) Transcriptional activation by recruitment. *Nature*, **386**, 569–577.

Pugh,B.F. (1996) Mechanisms of transcription complex assembly. *Curr. Opin. Cell Biol*., **8**, 303–311.

Roeder,R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci*., **21**, 327–335.

Shen,W.C. and Green,M.R. (1997) Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator. *Cell*, **90**, 615–624.

Shykind,B., Kim,J. and Sharp,P.A. (1995) Activation of the TFIID–TFIIA complex with HMG-2. *Genes Dev.*, **9**, 1354–1365.

Steinmetz,E.J. (1997) Pre-mRNA processing and the CTD of RNA polymerase II: the tail that wags the dog? *Cell*, **89**, 491–494.

Stelzer,G., Goppelt,A., Lottspeich,F. and Meisterernst,M. (1994) Repression of basal transcription by HMG2 is counteracted by TFIIH-associated factors in an ATP-dependent process. *Mol. Cell. Biol*., **14**, 4712–4721.

Tan,S., Conaway,R.C. and Conaway,J.W. (1995) Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc. Natl Acad. Sci. USA*, **92**, 6042–6046.

Triezenberg,S.J., Kingsbury,R.C. and McKnight,S.L. (1988) Functional dissection of VP16, the transactivator of herpes simplex virus immediate early gene expression. *Genes Dev.*, **2**, 718–729.

Verrijzer,C.P. and Tjian,R. (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci*., **21**, 338–342.

Verrijzer,C.P., Chen,J.L., Yokomori,K. and Tjian,R. (1995) Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell*, **81**, 1115–1125.

Walker,S.S., Reese,J.C., Apone,L.M. and Green,M.R. (1996) Transcription activation in cells lacking $TAF_{II}S$. *Nature*, **383**, 185–188.

Wang,W., Gralla,J.D. and Carey,M. (1992) The acidic activator GAL4-AH can stimulate polymerase II transcription by promoting assembly of a closed complex requiring TFIID and TFIIA. *Genes Dev.*, **6**, 1716–1727.

Zappavigna,V., Falciola,L., Citterich,M., Mavilio,F. and Bianchi,M.E. (1996) HMG-1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *EMBO J*., **15**, 4981–4991.

Zwilling,S., Konig,H. and Wirth,T. (1995) High mobility group protein 2 functionally interacts with POU domains of octamer transcription factors. *EMBO J*., **14**, 1198–1208.

Received July 27, 1998; revised and accepted September 15, 1998