Upstream stimulatory factor (USF) regulates a large variety of genes, including those for cell proliferation, differentiation, and metabolism. USF binds to E-boxes in DNA, forming homodimers or heterodimers with other transcription factors. The USF2 isoforms, which are produced by alternative pre-mRNA splicing, have been shown to interact with USF1 and USF2a to form active heterodimers.

**Upstream stimulatory factor (USF)**

The ubiquitous upstream stimulatory factor (USF) transcription factors encoded by two distinct genes (USF1 and USF2) exist under the form of various dimers able to bind E-boxes. We report the molecular cloning and functional characterization of USF2 isoforms, corresponding to a 44-kDa subunit, USF2a, and a new 38-kDa subunit, USF2b, generated by differential splicing. Using specific anti-USF antibodies, we define the different binding complexes in various nuclear extracts. In vivo, the USF1/USF2a heterodimer represents over 66% of the USF binding activity whereas the USF1 and USF2a homodimers represent less than 10%, which strongly suggests an in vivo preferential association in heterodimers. In particular, an USF1/USF2b heterodimer accounted for almost 15% of the USF species in some cells. The preferential heterodimerization of USF subunits was reproduced in vivo, while the in vitro association of cotranslated subunits, or recombinant USF proteins, appeared to be random. In transiently transfected cells, USF2a and USF1 homodimers transactivated a minimal promoter with similar efficiency, whereas USF2b, which lacks an internal 67-amino acid domain, was a poor transactivator. Additionally, USF2b was as efficient as USF1 and USF2a homodimers in transactivating the liver-specific pyruvate kinase gene promoter.

**Upstream Stimulatory Factor Isoforms**

**Benjamin Viollet, Anne-Marie Lefrancois-Martinez, Alexandra Hennion, Axel Kohn, Michel Raymondjean, and Antoine Martinez**

From the Institut Cochin de Genétique Moléculaire, U129 INSERM, Université René Descartes, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France.

*This work was supported by INSERM and grants from the Fondation pour la Recherche Médicale, the Association pour la Recherche sur le Cancer, the Ligue Nationale Française contre le Cancer, the Ministère de l’Enseignement Supérieur et de la Recherche, and the Institut Scientifique Roussel. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. The nucleotide sequence(s) reported in this paper has been submitted to the GenBank**/

**†‡To whom correspondence should be addressed. Tel.: 33-1-44-41-24-24; Fax: 33-1-44-41-24-21.**

**§Present address: Laboratoire de Reproduction et Développement, CNRS URA 1940, Université Blaise Pascal, Les Cézeaux, 63177 Aubière Cédex, France.**

**¶The abbreviations used are: USF, upstream stimulatory factor; MLP, major late promoter; L-PK, liver type pyruvate kinase; HLH, helix-loop-helix; b-HLH, basic HLH; bp, base pair(s); PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay; CAT, chloramphenicol acetyltransferase; CMV, cytomegalovirus; RT-PCR, reverse transcription PCR; USE, upstream stimulating element.**

© 1996 by The American Society for Biochemistry and Molecular Biology, Inc.
USF proteins family revealed by our present study. Therefore, we wondered about the respective role of the 43- and 44-kDa polypeptides, and eventually of other yet unidentified isoforms, in the USF activity involved in so diverse cellular functions. First, we report here the molecular cloning and functional characterization of the full-length human 44-kDa polypeptide USF2a and a new 38-kDa subunit USF2b. Second, in order to characterize the nature of the USF complexes in various cell lines and tissue extracts, we have developed monospecific polyclonal antibodies against different domains of the USF1 and USF2 proteins. These antibodies revealed to be powerful tools toward the definition of the various USF binding complexes.

We have found that USF1/USF2a heterodimer is predominant in all extracts but some variations between the relative proportion of each species is observed in different cells and tissues. In particular, more USF2b-containing species is detected in some erythroid cell extracts. Heterodimerization with USF proteins seems to be strictly restricted to USF family members; however, we have found that, in vivo, subunit association is biased toward heterodimerization, while a pattern of random dimer formation could only be observed in vitro. Finally, we have shown that, while USF1 and USF2a behave similarly on different promoters, USF2b potential transactivity could depend on the promoter context.

Study of Various USF Isoforms

MATERIALS AND METHODS
cDNA Libraries and Screening—Total RNA was isolated from tissues (human and rat liver) or from human cultured cells (HeP2G2, HeLa, lymphoblastoid Epstein-Barr virus-transformed and megakaryoblastic UT7 cells), using the guanidine isothiocyanate-cesium chloride method (37) and poly(A) RNA fraction was purified by oligo(dT)-cellulose chromatography (38). A rat liver cDNA library was constructed in λgt11 using random hexamers according to the manufacturer’s instructions (Amerham Corp.). Oligonucleotides located upstream ACGACTCGG-GATGGAAACG and downstream CAACGCTTGTTACTCTGCC from the nucleotide sequence encoding the β-hLH region of the human USF1 cDNA (nucleotides 700–903) were used to amplify a 204-bp fragment by PCR from pET3d USF43 as template (16). The random labeled probe was used to hybridize nylon filters (Hybond N, Amersham) at low stringency conditions. Sequencing of two resulting clones (L17 and L18) revealed that they corresponded to incomplete USF2 cDNAs slightly longer (L17) than those previously cloned in human (3, 19). A 267-bp probe, amplified by PCR, was designed in a USF2-specific region close to 5′ of L17 unique, using GCTGCCTTCGCAGGGGG and TCCGCGTCTAGGCTCTGCTG as upstream and downstream primers and was used to screen a human liver cDNA library cloned in in10 and primed with oligodT (Clontech). The filters were hybridized at high stringency conditions, and resulting λ clones were subcloned into the EcoRI site of the vector pBKS+ (Stratagene) and submittted to nucleotide sequencing on both DNA strands.

RT-PCR Analysis—cDNA synthesis was performed in standard conditions with 2 μg of total RNA, in presence of 0.75 mM each dNTP, 5 pm random hexanucleotide primer (Pharmacia Biotech Inc.), 20 units of RNasin (Promega), and 200 units of Moloney murine leukemia virus reverse transcriptase (Life Technologies, Inc.). PCR amplification was performed with 5 μl reverse transcriptase reaction mixture in standard conditions with 2.5 units of AmpliTaq polymerase (Perkin-Elmer). The amplified products were separated by electrophoresis on 6% polyacrylamide gels, blotted as described previously (39), and hybridized with both 150 bp USF2-specific PCR fragment and 243-bp β-actin fragment labeled by random priming.

Sequences of the oligonucleotides were as follows, for USF2a cloning:

USF2a: 5′-CAAGAGTCTGAGGACCTCGAGACCC
3′-GCATGTGTCCTCCCTCGT

and, for semiquantitative analyses of USF2 isoforms (in Fig. 4).

USF2a/USF2b: 5′-GCGATCTGGACGACCG
3′-TCCAGCTTGAACTGCTGTC

150-bp probe: 5′-GCTGTGATCCAAAAATCC (primer 2)

3′-TTGCCTTTAGGTTGCAAGGG

In Vitro Transcription/Translation and Protein Expression—The plasmids used for in vitro transcription were obtained by digesting the USF2 cDNA fragments (USF2b from p2 and USF2a from p9, see Fig. 1A) into the vector KTRAB (kindly provided by Dr. Y. De Keyzer). The plasmids containing mini-USF1 and mini-USF2 cDNAs used in transcription/translation assays were described elsewhere (13). Mini-USF2 cDNAs lacked NH2-terminal domains of USF proteins (USF1 Δ1–163 and USF2a Δ1–198, respectively). In vitro transcription of USF proteins was performed in rabbit reticulocyte lystate (Promega) using in vitro capped and transcribed RNA (Boehringer Mannheim).

The USF2a and USF2b cDNA were subcloned in the prokaryotic T7-based expression vector pET-28a (Novagen). Recombinant USF2a and USF2b proteins expressed in BL21(DE3)plyS cells were purified according to manufacturer’s instructions (Novagen). Recombinant USF1 protein expressed in BL21(DE3)plyS cells from pET3d USF43 vector was purified by ammonium sulfate precipitation (40).

Heterodimeric complexes between USF proteins were obtained by cotranslation of approximately equal amount of RNA templates or by mixing different amounts of recombinant proteins under denaturing conditions followed by an extensive dialysis.

Northern Blot Analysis—Extracts Preparation and Gel Shift Assays—Nuclear extracts were prepared from adult rat liver and fetal rat liver at day 17 of gestation according to Gorski et al. (41). Nuclear extracts or whole cell extracts from cultured cells (AT3F, HeP2G2, HeLa, and UT7 cells) were prepared according to Dignam et al. (42) or to Helin et al. (43).

Electrophoretic mobility shift assays (EMSA) were performed with either cell-free translation products, nuclear extracts, whole cell extracts or protein extracts used in CAT assays. The DNA binding reactions were performed as described previously (44) at 4°C in binding buffer (20 mM Hapes, pH 7.6, 50 mM KCl, 5 mM MgCl2, 4 mM spermine, 0.2 mM EDTA, 5 mM dithiotreitol, 4% (w/v) Ficoll) in the presence of 2–2 μg of nuclear extracts or 1–3 μl of reticulocyte lysate, 15 μg of whole cell extract and 2.5–10 ng of end-labeled USE/MLP oligonucleotide (45). For gel retardation assays with recombinant proteins, DNA binding reactions were performed in the presence of 20–40 ng of purified USF proteins, 100 ng of poly(dI-dC), 1 μl of preimmune rabbit antiserum, and 0.1–0.5 ng of end-labeled MLP oligonucleotide. For competition assays, 5 and 10 ng of USF/MPL probe were used as specific competitor and 10 ng of NF-Y probes unrelated competitor (44) and, for supershift assays, 1 μl of polyclonal antiseras was included in the binding reactions. The quantification of the DNA binding complexes was monitored using a PhosphorImager and analyzed using ImageQuant (Molecular Dynamics).

Antibody Production—Glutathione S-transferase fusion proteins were produced by subcloning specific USF1, USF2a, and USF2b fragment encoding expression vector pGEX-3X (Pharmacia). The genes encoding USF1-specific domain (amino acids 83–157, domain M), USF2a and USF2b common domain (amino acids 1–49, domain D) and the USF2a-specific domain (amino acids 93–143, domain O) were amplified by PCR with specific sets of oligonucleotides and cloned in-frame with the glutathione S-transferase moiety. Recombinant proteins were expressed in BL26 Blue cells and purified by glutathione affinity chromatography (Pharmacia). Fusion proteins were used to raise antibodies in rabbits according to a standard immunization procedure (Neosystem, Strasbourg, France).

Transfection and CAT Assays—To express the USF proteins in co-transfection experiments, we used the cytomegalovirus (CMV) immediate promoter to drive the expression of the USF1 and USF2 cDNAs. The USF cDNAs were inserted as blunt-ended fragments in the filled-in EcoRI site of the pCMV vector (kindly provided by Dr. R. H. Costa). CAT reporter plasmids, −54PKCAT, −183PKCAT, and four tandemly repeated USE/MLP or L4 fused to −54PKCAT have been described previously (13, 46, 47). Hela cells were cultured in Dulbecco’s modified Eagle’s medium containing 5% fetal calf serum and, as described above with the same protein extracts as those used in CAT assays. CAT activities were calculated as the percentage of acetylated form of chloramphenicol versus the non-acetylated substrate. The ac-
activities of the different USF isoforms were defined as the highest CAT activity correlated with the ectopic protein produced. Each experiment was made in duplicate and repeated two to four times.

RESULTS

Isolation of a Family of Alternatively Spliced Human USF2 cDNAs—To isolate complete USF2 transcription factors cDNAs, a λgt11 random primed rat liver cDNA library was screened at low stringency conditions with a probe overlapping the b-HLH region of the human 43-kDa USF1 cDNA. From 500,000 recombinants, we isolated two phage clones. The larger open reading frame encoded a polypeptide highly related to human USF1 43-kDa protein only in its COOH-terminal part (67% identity) and almost identical to the previously reported human USF2 cDNAs (14, 19). A 267-bp probe overlapping the USF2-specific region of the rat cDNA was used in turn to screen, at high stringency conditions, a λgt10 human liver cDNA library. This screen resulted in the isolation of 13 positive plaques from 800,000 recombinants.

As outlined in Fig. 1A, the clone p9ΔH appeared still as a partial cDNA lacking an initiating methionine codon. In contrast, p2 clone was probably a full-length cDNA, containing a 5′ untranslated region (109 bp) and an open reading frame of 279 amino acids starting at a putative ATG initiation codon. Several lines of evidence indicate that the p9ΔH and p2 clones correspond to two different transcripts from the same gene, generated by differential splicing. (i) p9ΔH and p2 clones differed from each other only by their 5′ ends and by the length of the poly(A) tail. (ii) clone p9ΔH included an internal 201-bp sequence (67 amino acids) that was absent from p2. From 12 RT-PCR products amplified from human liver mRNA using two primers located at the two ends of the p2 clone, four were found identical to p2 and eight contained the expected complete open reading frame of 346 residues attributed to the complete p9 sequence (see Fig. 1A). Thus, p2 differs from p9 by the alternative splicing out of 201 bp (boxed in Fig. 1B). (iii) examination of clone p9ΔH sequence pointed out a second splicing
pattern. Indeed, a 24-bp fragment encoding the first 8 residues of the second helix of the HLH domain was missing. Intriguingly, the remaining reading frame was not modified so that the putative polypeptide exhibited a partially deleted HLH domain but did contain an intact leucine zipper domain. This deleted transcript is naturally generated since RT-PCR strategy allowed us to amplify helix 2-deleted cDNAs, using sense oligonucleotide primers overlapping the deletion point and antisense primers derived from the 3’ end of the transcribed region. However, this isoform was very poorly represented in human cells and was virtually undetectable in rat cells (data not shown). We termed the cDNA clones p9Ah, p9, and p2, respectively, hUSF2aAh, hUSF2a, and hUSF2b (right side in Fig. 1, A and B). Finally, sequence analysis of genomic clones of murine USF2 gene indicates that the optional 201-bp region missing in USF2b cDNA lay within the fourth exon (5, 20), while the Ah isoform is generated by an additional splicing event within the ninth exon, i.e., the use of a cryptic splicing acceptor site located 24 bp downstream from the normal splice site (5). The in vitro translated protein hUSF2aAh was unable to bind DNA in band shift assay, probably reflecting its incapacity to dimerize (data not shown).

Immunochemical Characterization of the Different USF Species in Liver Nuclear Extract—As a first step toward identifying the protein components of oligomeric USF complexes present in nuclear extracts, we developed specific antibodies against USF1 and USF2 proteins. Polyclonal antibodies were raised in rabbits against nonconserved amino acid domains (Fig. 2A). Domain M of USF1 protein encompasses amino acids 83–157, domains G and O of USF2 proteins, amino acids 1–49 and 93–143, respectively. The latter domain, corresponding to the fourth exon of USF2 gene (5, 20), allows one to distinguish USF2a from the USF2b isoform. The antibodies raised against these motifs are designated IgG M, IgG G, and IgG O. The serum termed IgG 4, developed by Pognonec and Roeder (49), was directed against amino acids 273–310 of USF1 protein including 3 of the 4 leucine residues of the leucine repeat; similarly, IgG Z was raised against amino acids 298–346 of USF2a/2b protein corresponding to the entire leucine zipper motif. The specificity of these antibodies was tested by EMSA with recombinant or in vitro cotranslated proteins, resulting in supershifted USF-DNA complexes that remain at the top of the gel (data not shown and Fig. 5).

Therefore, our different antibodies appeared to be suitable reagents for identifying the nature of the different USF complexes detected in various cells and tissues. Fig. 2B shows that the major complex detected in liver nuclear extract was displaced by both anti-USF1 (IgG M) and anti-USF2 (IgG G and O) antibodies, demonstrating that it corresponds to a USF1/USF2a heterodimer (compare lanes 4–6 to lane 7). After having supershifted this heterodimer, three faint bands, masked by the major species, could be clearly resolved. Indeed, the slow migrating band was displaced by anti-USF 2 (IgG G and IgG O) but not anti-USF 1 (IgG M) antibodies and could be identified as USF2a homodimer (lane 6). The faster migrating, USF1 homodimer, displaced by IgG M but neither IgG G nor O, was also revealed (lane 5). Finally, as the very faint band slightly below the USF1 homodimer was displaced by IgG M and IgG G, but not IgG O, it could be identified as USF1/USF2b heterodimer (lane 4). These identifications were consistent with the electrophoretic mobility of these different species compared to that of the various dimeric forms resulting from in vitro cotranslation of either USF2a and USF2b mRNAs (lane 2) or USF1 and USF2b mRNAs (lane 3), as shown by the position of USF1 and USF2a homodimers and USF1/USF2b heterodimers indicated on the left of Fig. 2B.

The exact composition of the two faster migrating complexes A and B was not clear. They could be considered as heterodimeric complexes between USF3 (complex B) or USF2a (complex A) full-length subunits and shorter partners. The supershifted complexes with anti-USF 1 ZIP antibodies, IgG 4 (lane 13), and anti-USF 2 ZIP antibodies, IgG Z (lane 15), and the comigration with in vitro cotranslated proteins complexes (lanes 8–11) indicated the presence in these complexes of NH2 truncated USF proteins. These latter USF proteins were described as mini-USF (15). In vitro treatments of full-length USFs by the calf pancreatase, which cleaves the NH2-terminal ends, gave rise to faster migrating forms (data not shown). This result strongly suggests that mini-USFs are translated by the use of internal methionines. Finally, complex A can be defined as a USF2a/mini-USF1 heterodimer (compare lanes 6 and 9) and B as a USF1/mini-USF2 heterodimer (compare lanes 5 and 10).

Quantification of the Different USF Dimers Detected in Various Cells and Tissues—The nature and relative abundance of the different USF complexes in various nuclear extracts were investigated by gel shift assays (Fig. 3B). The specificity of different retarded complexes was checked by competition with unlabelled oligonucleotides in excess: either specific (MLP) or unrelated (NF-Y) (Fig. 3A). The bands marked by an asterisk are nonspecific since they were not competed for by an excess of MLP oligonucleotide. Quality of the extracts was systematically assessed by analyzing the ubiquitous NF-Y and NF-1 binding activities; indeed, these transcriptional factors are known to be very sensitive to proteolytic cleavages (44) (data not shown). USF1 and USF2a homodimers were quantified by measuring the band intensities observed after respective incubation with antibodies IgG M and IgG G. USF1/USF2b heterodimers were quantified from the band migrating slightly below USF1 homodimers after treatment with antibodies IgG O. Table I shows that USF1/USF2a heterodimers are predominant in all tested cells and tissues, ranging from 66 to 76% of the total USF binding activity. The USF1 and USF2a homodimers appeared to be minor complexes, ranging from 5 to 10% of the USF binding activity in most extracts. USF1 homodimer reached 18% of the USF binding activity in HeLa cells. USF1/USF2b heterodimer was especially represented in fetal liver (erythropoietic organ) and erythroid UT7 cells, where it reached up to 15% of the USF binding activity. An identical pattern was observed using nuclear extracts from other erythroid-specific cell lines, Red 707 and MEL J 3 cells (data not shown). In these nuclear extracts, the increase of USF1/USF2b heterodimers was associated with a decrease of USF2a homodimers. Such an increase of the USF2b versus USF2a isoforms in erythroid cells may reflect a modulation of the ratio of USF2 transcripts spliced into USF2b.

To test this hypothesis, we used a semiquantitative RT-PCR strategy illustrated in Fig. 4A. Analysis of PCR products on cDNAs from different cell lines and tissues showed the two predicted USF2 fragments of similar intensities (Fig. 4B). Northern blot analysis suggested that the USF2 gene was expressed at a similar level in all tested cells (data not shown), even though the ratio between USF2a and USF2b-specific amplified fragments varied significantly. Indeed, USF2b mature messenger seems to be more accumulated in HeLa and UT7 cells than in adult liver, hepatoma HepG2, and lymphoblastic cells. From these data, we infer that different cellular contents in USF1/USF2b complex might occur from a regulation of alternative splicing of USF2 precursor mRNA.

In order to quantify the different USF dimers according to the gel shift patterns, we checked that these complexes exhibited similar binding activity. We therefore compared the pat-
terns using different oligonucleotide probes of decreasing affinity for USF proteins (data not presented). All complexes generated with either crude liver extracts or the three recombinant USF proteins were observed to display similar affinities for each of the probes. Thus, these results indicate that USF homo- and heterodimers shared similar DNA binding properties in vitro, and that binding activities of the various heterodimeric isoforms present in crude nuclear extracts are not significantly different.

Heterodimeric Preferential Association of USF Proteins in Vivo—It is quite interesting to note the prevalence of heterodimeric complexes in the different nuclear extracts analyzed. The most striking example is the high proportion of USF1-USF2 heterodimers compared to the USF1 and USF2a homodimers (Table I). In erythroid cell lines where USF2b expression is increased, USF1-USF2b heterodimers were readily visualized in extracts treated with antibodies IgG O (Fig. 3B), whereas USF2a/USF2b heterodimers and USF2b homodimer were barely undetectable. As heterodimeric association seems to be preferred in vivo...
according to our results, we then tested this preference in vitro by using recombinant USF and cell-free translated proteins in EMSA. When recombinant USF proteins were mixed, dissociated under denaturant conditions, and then reassociated by removal of the denaturant, we observed that the USF1/USF2 heterodimers complexes were formed at random (Fig. 5A). A 1:2:1 ratio was observed for the different complexes formed. When USF proteins were cotranslated in vitro using equimolar ratios of USF1 and USF2a mRNAs, we still obtained a random dimerization pattern (Fig. 5B). The use of specific antibodies allowed us to discriminate homodimers from heterodimers of intermediate electrophoretic mobility. Furthermore, we did not observe any subunit exchange incubating homodimeric recombinant USF proteins in ionic strength conditions performed in nuclear extract preparation (0.42 M NaCl extraction and 50% (NH4)2SO4 precipitation). These results indicate that homodimerization and heterodimerization of the USF1 and USF2 isoforms are similar in vitro, and suggest that preferential association be-

**Fig. 3.** Analysis of cell-specific USF DNA binding patterns. A, nuclear extracts were from rat tissues (adult liver and fetal liver), human cells (hepatoma cells (HepG2), megakaryoblastic cells (UT7), fibroblasts (HeLa)), and mouse hepatoma cells (AT3F). EMSA were carried out by using the radiolabeled MLP oligonucleotide. Five and 10 ng of specific unlabeled oligonucleotide MLP, and 10 ng of unrelated oligonucleotide NF-Y (44), were added in competition experiments (lanes MLP and NF-Y). Alongside the figure are indicated the position of the heterodimeric USF1/USF2a and USF1/USF2b complexes. A and B correspond to USF2a/U1 and USF1/U2 heterodimers, respectively. B, representation of the various USF DNA binding complexes in different tissues and cells. EMSA was performed as described above and selective depletion of the complexes was realized by adding specific anti-USF antibodies (IgG) to the binding reactions. IgG M is specific to USF1 protein, IgG G specific to USF2a and USF2b proteins, and IgG O specific to USF2a protein. Lanes (–), naive serum. Complexes revealed are mentioned alongside the figure. The asterisks indicate nonspecific protein complexes.
between USF1 and USF2 gene products is specific feature of in vivo conditions (Table I).

To confirm this hypothesis, fixed amounts of USF2a or USF1 expression vectors were cotransfected in HeLa cells with increasing amounts of USF2b. Oligomerization properties of the ectopic USF proteins were analyzed by EMSA, depleting the whole cell extracts from endogenous USFs with anti-USF1 or anti-USF2a antibodies (Fig. 5C), and the different protein complexes visualized were quantified. When ectopic USF2a versus USF1, and USF1 versus USF2b were produced in comparable amounts (boxed in Fig. 5C), the dimer distribution displayed very different patterns. Indeed, whereas USF2a and USF2b were able to associate in a 1:2:1 distribution, which could be explained considering that each monomer had no preference for a dimerization partner, USF1/USF2b heterodimeric association was clearly favored over homodimerization. It appears from these results that in vivo as well as in vitro, all USF family members are able to form heterodimers with each other. However, in vivo but not in vitro, USF2 isoforms preferentially heterodimerized with USF1.

Transactivation Activities of the Different USF Isoforms—The ability of USF2 alternative forms to function as transcriptional regulators was analyzed by transient transfection in HeLa cells and compared with the USF1 transactivating efficiency (50). We constructed three different CAT reporter plasmids driven by the (MLP)4–54PKCAT with or without one or four tandem repeats of the MLP site placed upstream from the L-PK minimal promoter (L-PK) (46) with reporter genes containing MLP sites (MLP-54PKCAT with USF1 or USF2b). Oligomerization properties of the ectopic USF proteins were anayed by EMSA, depleting the whole cell extracts from endogenous USFs with anti-USF1 or anti-USF2a antibodies (Fig. 5C), and the different protein complexes visualized were quantified. When ectopic USF2a versus USF1, and USF1 versus USF2b were produced in comparable amounts (boxed in Fig. 5C), the dimer distribution displayed very different patterns. Indeed, whereas USF2a and USF2b were able to associate in a 1:2:1 distribution, which could be explained considering that each monomer had no preference for a dimerization partner, USF1/USF2b heterodimeric association was clearly favored over homodimerization. It appears from these results that in vivo as well as in vitro, all USF family members are able to form heterodimers with each other. However, in vivo but not in vitro, USF2 isoforms preferentially heterodimerized with USF1.

Table I

| USF complexes | Adult liver | AT3F | HepG2 | Fetal liver | UT7 | HeLa |
|---------------|-------------|------|-------|-------------|-----|------|
| U1/U2a        | 66          | 73   | 76    | 75          | 67  | 67   |
| U2a/U2a       | 8.6         | 5.2  | 8.7   | 6           | 8   | 6.6  |
| U1/U1         | 7.2         | 6.4  | 7.3   | 10.8        | 10  | 17.6 |
| U1/U2b        | 3.6         | 4.5  | 7.6   | 8           | 15  | 9    |
| U2a/U1        | 8.3         | 6.4  | ND    | ND          | ND  | ND   |
| U1/U2a        | 5.8         | 4.1  | ND    | ND          | ND  | ND   |

Table I shows a comparison of USF binding activities in various nuclear extracts. DNA binding activity of the various USF complexes is given as a percentage of overall binding activity. ND, nondetectable complexes.

Fig. 4. Cell-type variations of USF2 message splicing. A, RT-PCR were conducted with total RNAs from human cells using primers 1 and 3. The probe fragment used to hybridize the subsequent blot was amplified using primers 2 and 3 and the plasmid p9 (hUSF2a) as template. To ensure semiquantitative analysis, the amplification process was maintained in exponential phase. Respective sizes of predicted RT-PCR products are indicated. B, amplified fragments were electrophoresed in a polyacrylamide gel. The reaction specificity was checked using template plasmids p9 and p2 (containing hUSF2a and hUSF2b cDNAs, respectively) as positive control and a mixture of all nonreverse-transcribed cellular RNAs as a negative control (ARN mix). Control of the amount of total cDNA present in each lane was checked using β-actin-specific primers. The gel was blotted and hybridized with the common USF2 150-bp probe and a β-actin probe. Total RNAs were from rat liver (Li), human HepG2 hepatoma cells (Hep), HeLa cells (HeLa), human megakaryoblastic cells (UT7), and human lymphoblastic Epstein-Barr virus-transformed cells (Ly).

As shown at the top of Fig. 6C, cotransfections of (MLP)4–54PKCAT with USF1 or USF2a expression vectors resulted in a very efficient stimulation of CAT activity while USF2b plasmid stimulated poorly. As estimated from EMSA using the same extract than in the CAT assay, the extent of activation essentially depended on the amount of protein produced (Fig. 6C, bottom). The maximum of stimulation occurred at 2, 0.5, and 1 μg of USF1, USF2a, and USF2b expression vectors transfected, respectively. Under these conditions, USF1, USF2a, and USF2b, respectively, resulted in 23-, 18-, and 6-fold transcription stimulation relative to the endogenous USF activity revealed by the empty expression vector. We conclude from these experiments that USF2 gene products have different effects on transcription modulation through the USE-MLP site. USF2a is a transcription activator as efficient as USF1, while the USF2b isoform is at least 3 times less active under the same experimental conditions. This result indicates that the 67-amino acid region of USF2a, missing from USF2b, may be involved in the transcriptional activation.

Next, we compared the respective ability of USF isoforms to modulate the activity of the L-PK promoter that has been shown previously to be dietary controlled via a USF-dependent glucose responsive element, GIRE, consisting in two E-boxes (element L4) (11–13, 47). The trans-modulating activity mediated by each USF expression plasmid was tested by transient
cotransfections in mouse hepatoma AT3F cells using a CAT reporter gene driven by the L-PK promoter (Fig. 7A). Surprisingly, as shown in Fig. 7B (left side), the ectopic production of all USF isoforms even USF2b resulted in a similar 4–5-fold stimulation of the CAT activity relative to the endogenous activity observed with the empty expression vector. To know whether this unexpected USF2b activity was dependent on the L-PK promoter context or might be assigned to the L4 element itself, we performed cotransfections with a reporter gene carrying four copies of the L4 element placed upstream from the MLP minimal promoter ((L4)4–54PK, Fig. 7A). As shown in Fig. 7B (right side), overexpression of USF1 and USF2a stimulated transcription about 60-fold while USF2b resulted in a 18-fold stimulation relative to the endogenous activity for equivalent DNA binding activity checked by EMSA. These results obtained with the (MLP)4–54PK CAT plasmid were cotransfected with various amounts of either the empty or the USF-containing expression vectors as indicated (in μg) under the picture. All values represent the means of two or three independent experiments in duplicate. EMSA were performed with the MLP probe and the same whole cell protein extracts used in CAT assays. Anti-USF1 (IgG M) and anti-USF2 (IgG O) antibodies were added in reaction mixtures to deplete the extracts of endogenous USF1/USF2 binding activity. The asterisk indicates nonspecific protein complex.

Fig. 6. Analysis of the USF1 and USF2 isoforms transacting potential in HeLa cells. A, schematic representation of the CAT reporter plasmids containing various repeats of the USE-MLP site upstream from the L-type pyruvate kinase minimal promoter and of the USF expression plasmids driven by the CMV promoter. B, transient transfection assays in HeLa cells of the reporter CAT constructs in presence of USF1 and USF2a expression vectors. HeLa cells were cotransfected by calcium phosphate procedure with 5 μg of indicated CAT reporter plasmids and 2 or 0.5 μg of expression plasmids encoding, respectively, USF1 and USF2a (pCMV-U1 and pCMV-U2a) or 2 μg of control plasmid pCMV. C, the transacting effects of the USF1 and USF2 expression vectors are compared. 5 μg of the (MLP)4–54PK CAT plasmid were cotransfected with various amounts of either the empty or the USF-containing expression vectors as indicated (in μg) under the picture. All values represent the means of two or three independent experiments in duplicate. EMSA were performed with the MLP probe and the same whole cell protein extracts used in CAT assays. Anti-USF1 (IgG M) and anti-USF2 (IgG O) antibodies were added in reaction mixtures to deplete the extracts of endogenous USF1/USF2 binding activity. The asterisk indicates nonspecific protein complex.

Fig. 5. Oligomerization properties of USF1 and USF2 isoforms in vitro and in vivo. A, recombinant USF proteins were mixed in various ratio under denaturing conditions and the complexes formed after removal of the denaturant by dialysis, were visualized by EMSA using the MLP radiolabeled probe. B, in vitro cotranslated USF protein complexes were analyzed by EMSA using the MLP-radioabeled probe. Cotranslation products were incubated with anti-USF antibodies to discriminate homodimeric and heterodimeric USF complexes. C, EMSA analysis of ectopic USF complexes binding to MLP probe in HeLa cells cotransfected with fixed amounts of pCMV-U2a (4 μg) and pCMV-U1 (6 μg) and increasing amounts of pCMV-U2b. Antibodies anti-USF1, IgG M (pCMV-U2a/U2b cotransfections), or anti-USF2a, IgG O (pCMV-U2a/U2b cotransfections), were systematically added in reaction mixtures to deplete the whole cell extracts of endogenous USF binding activity. The percentage of each ectopic USF complex (arrows) is indicated. The percentages corresponding to approximately equal expression of each partner are boxed. Excess of unlabeled MLP oligonucleotide was used in a competition assay in lane C. The asterisk indicates nonspecific protein complex.
USF DNA binding activity in HeLa cells was initially described as a complex consisting of two polypeptides of 43 and 44 kDa (4). Various alternatively spliced USF transcripts were previously described resulting in non-coding RNAs (14–16). In this paper, an additional level of complexity is described since we show that the human USF2 gene generates, by differential splicing, different mRNAs encoding polypeptides with respective apparent molecular masses of 44 kDa (USF2a) and 38 kDa (USF2b). These two isoforms share perfect identity in an extended COOH-terminal region of 203 amino acids, encompassing the b-HLH-ZIP motifs, but vary in the organization of their NH₂-terminal domains (Fig. 2A). Indeed, USF2b lacks an internal 67-amino acid domain present in USF2a. Analysis of murine genomic clones showed that the USF2b messenger occurs by the splicing out of the fourth exon in primary transcripts (5, 20). In addition, USF2aH, an additional spliced form, although containing an intact leucine zipper domain, lacks the second helix of the HLH motif due to a 24-bp in-frame deletion occurring by the use of a cryptic splicing site within the ninth exon (5). As suggested by RT-PCR analysis (data not shown), this latest splicing event is poorly efficient in human and virtually undetectable in rat cells. The biological relevance of the JH isoform also remains unclear.

As an effort to characterize the various USF DNA-binding complexes, we have developed various polyclonal antibodies against different regions of the USF1 and USF2 proteins. Our immunohistochemical analyses in nuclear extracts from various cell lines and tissues revealed that the major USF species detected was a USF1-USF2a heterodimer, corresponding to the species composed of the 43- and 44-kDa subunits. In addition, minor forms whose abundance varies in different cell types and tissues could be detected. A heterodimer between USF1 and USF2b subunits was especially abundant in some cultured erythroid cell lines, where it reached 15% of the total binding activity. Consistent with ubiquitous distribution, RT-PCR and EMSA analyses showed that USF2a and USF2 isoforms are coexpressed in every cell types tested with slight cell type preferences concerning USF2b expression. Indeed, we observed a modulation in the amount of USF1-USF2b binding complex, which is well correlated to the relative USF2b mRNAs accumulation observed by RT-PCR. Interestingly, the greater expression of USF2b isoform in UT7 cells and fetal liver seems to be a common feature of the erythroid cell type. It was previously observed that relative abundance of USF1 (43 kDa) and USF2 (44 kDa) homo- and heterodimers seemed to vary among different cell types (15). We report here the presence of a novel USF binding complex occurring by a quantitative modulation of the USF2 pre-mRNA splicing. Whether this modulation corresponds to possible tissue-specific regulatory functions remains to be carefully investigated. Two low molecular weight complexes were also detected in liver-specific cells, i.e. nuclear extracts from liver and cultured mouse hepatoma AT3F cells, which could be identified as heterodimers between full-length USF1 or USF2a proteins on the one hand, mini-USF2a or mini-USF1 on the other. Mini-USF proteins are deleted from their NH₂ termini containing the transactivation domains and are likely to result from the use of internal translation initiators as proposed by Sirito et al. (15). Indeed, we observed that similar NH₂-truncated forms resulting from calpain digestion (56) migrated faster than these mini-USFs (data not shown).

In rat liver nuclear extracts as in all cells tested, USF binding activity corresponds mainly to four different complexes as presented by our gel shift experiments. In vivo, most USF species are heterodimers between USF1 and USF2a isoforms, with an obvious under-representation of homodimers. The fact that USF1 and USF2 gene products preferentially het-

Fig. 7. Transacting effects of USF1 and USF2 isoforms on the glucose-responsive element (GIRE) of the L-type pyruvate kinase promoter in mouse AT3F hepatoma cells cultured without glucose. A, schematic representation of the CAT reporter plasmids containing either the −183 L-PK promoter or four copies of the L4 element located upstream from the L-type pyruvate kinase minimal promoter, and of the USF expression plasmids driven by the CMV promoter. Boxes L1–L3 bind HNF1, NF1, and HNF4 proteins respectively (45). B, transient transfection assays in AT3F cells of the reporter CAT constructs in the presence of USF expression vectors. Comparison between the transacting effects of the USF1 and USF2 expression vectors. 5 μg of the −183PKCAT or of the (L4)4–54PKCAT plasmids were cotransfected with indicated amounts of either the empty or the USF-containing expression vectors. All values represent the means of at least three independent experiments in duplicate. EMSA were performed with the MLP probe and the same whole cell protein extracts used in CAT assays. Anti-USF1 (IgG M) and anti-USF2 (IgG G) antibodies were added in reaction mixtures to deplete the extracts of endogenous USF1/USF2abindingactivity. The asterisk indicates nonspecific protein complex.

USF proteins are ubiquitous transcription factors that were initially considered playing a probable role in housekeeping functions (51). Furthermore, these USF proteins have also been recognized as important players in the regulation of tissue-specific genes (52–55) and recently in the specific response of genes to external modulators (10), as glucose (11–13). So far, how such a type of ubiquitous factor can be involved in highly regulated transcriptional phenomena remains unknown. Therefore, it seems of prime interest to be able to accurately identify and quantify the various USF isoforms in different tissues or cells, and to determine their transacting potential.
erodimerize was quite surprising. In fact, the specificity of association between USF subunits is probably allowed by the dimerization domain itself, which should then dictate the compatibility between the different family members as this is observed with the b-ZIP protein Fos, unable to dimerize with itself, while it forms stable dimers with Jun (57). The dimerization interfaces of the coiled-coil motif is supposed to play a critical role in the dimerization specificity of b-HL-ZIP proteins (33). Interhelical electrostatic interactions between opposite charged residues located in the leucine zippers of b-ZIP and b-HL-ZIP factors are needed to design specific dimerization partners (57–59). However, this mechanism seems unlikely to explain USF1/USF2 preferential dimerization since USF1 and USF2 species cotranslated in a cell-free system or associated after denaturation of recombinant proteins, dimerized randomly to give the exact 1:2:1 dimer distribution. These results indicate that the HLF-ZIP dimerization domains of USF proteins exhibit no particular preference for partners in vitro. However, the in vivo preferential association may be experimentally reconstituted and visualized in EMSA using HeLa whole cell extracts with high levels of USF1 and USF2 obtained by cotransfection, indicating that this phenomenon is strictly dependent on a cellular context. Therefore, the reason why the in vivo association pattern is biased toward USF1/USF2 heterodimers remains rather obscure. A preferential stability of heterodimers in the intracellular conditions could be a first explanation. In any case, such a stabilization could not be ascribed to an intrinsic higher affinity of the heterodimers for DNA since, at least in vitro, homo- and heterodimers appeared to have the same affinity for the various types of USF binding sites we tested (data not shown). In vivo preferential isoform associations might also be governed by post-translational control mechanisms. Recently, the redox state of b-HL-ZIP proteins has been shown to control dimerization and DNA-binding of E2A proteins (60). A redox control of transcription factor activity has also been demonstrated for Fos and Jun (61). In the case of USF1 protein, non-reducing conditions were shown to decrease the affinity for DNA recognition motifs (62).

According to earlier in vitro and in vivo studies, the 43-kDa USF1 gene product was supposed to account for all of the transcription stimulation activity present in the 43/44-kDa USF complex purified from Hela cells (49, 50). Using transfection assays, we demonstrated that the 44-kDa USF2a component is able to activate transcription of a reporter gene driven by either minimal promoters containing four copies of USF binding sites or by a natural USF-dependent promoter, e.g. the −183 L-PK promoter (13), as efficiently as USF1. However, the 38-kDa USF2b alternative spliced form revealed a promoter-dependent activity. Indeed, in experiments using minimal promoters (MLP)4–54PK or (L4)4–54PK, we showed that USF2b was a poor transactivator, 3–4 times less active through identical binding sites than USF1 or USF2a, while it stimulated −183 L-PK promoter as efficiently as USF1 or USF2a homodimers. In experiments using minimal promoters, the transactivating potential of the three USF variants was not cell type-dependent (HeLa cells or hepatoma cells) and occurred through multimerized high affinity binding sites (MLP-USE of adenovirus, core sequence: CACGGGCGATCTCCGGT). In agreement with a modular composition of activation domains among USF proteins, previous deletion experiments had allowed the identification of two activation domains (one NH2-terminal spanning residues 26–39 and a second spanning residues 105–130) that cooperate in vitro for full transcriptional activity of USF1 (63). Consistent with its stronger activation potential, the NH2-terminal domain was particularly well conserved among vertebrate USFs (70% homology). The second is not conserved between USF1 and USF2 proteins, suggesting that the USF1 and USF2 gene products could interact with different accessory factors or components of the transcription machinery. Although further investigations are required to identify USF2 transactivation domains, our findings suggest that the activation potential of USF2a requires the 67-amino acid insert deleted in USF2b. Noticeably, the transactivating potential of this alternatively spliced domain of USF2a seems to depend upon the promoter context. Indeed, its exclusion from USF2a, resulting in a USF2b protein that conserves the putative NH2-terminal transactivating domain, leads to a strong reduction of the transactivation potential through minimal promoters ((MLP)4–54PK or (L4)4–54PK) but has little or no effect in the L-PK promoter context. The involvement of multiple domains that cooperate to allow fine-tuning of transactivating ability, has been found in several transcription factors. For instance, three activator elements (TE1–TE11) are required for full C/EBPα transcriptional activity and one of them, TE11, is shown to contain a negative subdomain, the function of which is controlled depending on the promoter context (64). An alternative explanation may be that the transactivation potential of USF2b isoform mediated through the L4 motif (18-fold stimulation of (L4)4–54PK construct, Fig. 7B) should be enough to allow the full stimulation of the −183 L-PK promoter (maximum 5-fold).

There is increasing evidence that ubiquitously expressed factors can also play an important role in tissue-restricted expression as well as in highly controlled cellular functions, as for instance hormonal and dietary regulations (11–13). The mechanisms by which USF proteins contribute to these specific functions are yet unclear and may overlay several phenomena. The NH2 activation domains of USF1 and USF2a are different and therefore could interact with different transcription factors bound at proximity, auxiliary factors, co-activators, or general transcription factors (1, 51, 65–67). Therefore, the various USF dimers could have slightly different specificities for a number of genes. In particular, diverse proportions of activatory and inhibitory USF dimers in different tissues and stages of development could allow a fine tuning of gene expression during differentiation and transcriptional response to external effectors. Moreover, a high proportion of USF2b-containing dimers could result in a down-regulation of USF transactivating activity in some tissues. Since mini-USF proteins behave as transdominant inhibitors (13), their accumulation at a sufficient level in some cells could also be expected to down-regulate USF activity. However, although the abundance of USF2b and mini-USF species depends clearly on the cell types, whether their involvement in total USF activity is sufficient to play this putative negative regulatory role remains to be studied. The structural and functional diversities of the USF family, here defined, should enlighten the way of investigating the physiological roles of USF proteins.

Acknowledgments—We thank Drs. R. G. Roeder for the gift of the pET3dUSF43 plasmid, P. Pognonec for providing the anti-USF1 antibody, Y. De Keyser for the KTRAB plasmid, C. Lacombe for UT7 cells, and F. P. Thierry for HeLa cells. We are grateful to Drs. Y. De Keyser, P. Maire, M. Vasseur-Cognet, B. Wasylyk, and M. Yaniv for critical reading of the manuscript. We thank also P. Chafey for expert technical assistance and to P. Marienneau and J.-E. Guidotti for helpful discussions and advice.

REFERENCES
1. Sawadogo, M., and Roeder, R. G. (1985) Cell 43, 165–175
2. Chodosh, L. A., Carthew, R. W., and Sharp, P. A. (1986) Mol. Cell. Biol. 6, 4723–4733
3. Sawadogo, M., Van Dyke, M. W., Gregor, P. D., and Roeder, R. G. (1988) J. Biol. Chem. 263, 11985–11993
