Upstream stimulatory factor activates the vasopressin promoter via multiple motifs, including a non-canonical E-box

Judy M. COULSON*†‡[1], Jodie L. EDGSON‡, Zoe V. MARSHALL-JONES*†, Robert MULGREW‡, John P. QUINN*† and Penella J. WOLL‡[2]

*Physiological Laboratory, University of Liverpool, Crown Street, Liverpool L69 3BX, U.K., †Department of Human Anatomy and Cell Biology, University of Liverpool, Ashton Street, Liverpool L69 3GE, U.K., and ‡Cancer Research UK Department of Clinical Oncology, University of Nottingham, City Hospital, Hucknall Road, Nottingham NG5 1PB, U.K.

We have described previously a complex E-box enhancer (−147) of the vasopressin promoter in small-cell lung cancer (SCLC) extracts [Coulson, Fiskerstrand, Woll and Quinn, (1999) Biochem. J. **344**, 961–970]. Upstream stimulatory factor (USF) heterodimers were one of the complexes binding to this site *in vitro*. We now report that USF overexpression in non-SCLC (NSCLC) cells can functionally activate vasopressin promoter-driven reporters that are otherwise inactive in this type of lung cancer cell. Site-directed mutagenesis and electrophoretic mobility-shift analysis demonstrate that although the −147 E-box contributes, none of the previously predicted E-boxes (−147, −135, −34) wholly account for this USF-mediated activation in NSCLC. 5′ Deletion showed the key promoter region as −52 to +42; however, USF-2 binding was not reliant on the −34 E-box, but on a novel adjacent CACGGG non-canonical E-box at −42 (motif E). This mediated USF binding in both SCLC and USF-2-transfected NSCLC cells. Mutation of motif E or the non-canonical TATA box abolished activity, implying both are required for transcriptional initiation on overexpression of USF-2. Co-transfected dominant negative USF confirmed that binding was required through motif E for function, but that the classical activation domain of USF was not essential. USF-2 bound motif E with 10-fold lower affinity than the −147 E-box. In NSCLC, endogenous USF-2 expression is low, and this basal level appears to be insufficient to activate transcription of arginine vasopressin (AVP). In summary, we have demonstrated a novel mechanism for USF activation, which contributes to differential vasopressin expression in lung cancer.

Key words: arginine vasopressin (AVP), enhancer, initiation, lung cancer, upstream stimulatory factor (USF).

INTRODUCTION

Small-cell lung cancer (SCLC) is characterized by the production of many neuropeptides that can act as growth factors for these tumours [1,2]. While some peptides are common to all types of lung cancer [3], other neuropeptides such as arginine vasopressin (AVP) are more highly restricted [4,5] and are not normally expressed in bronchial epithelial cells or non-neuroendocrine tumours, including the majority of non-SCLCs (NSCLCs) [6]. Physiological vasopressin expression is largely restricted to the hypothalamus and is tightly regulated by plasma osmolality [7]. In contrast, the ectopic expression of vasopressin in SCLC appears to be uncoupled from these controls. Our previous data would indicate that upstream stimulatory factor (USF) might play a role in this process. We have used non-neuroendocrine NSCLC as a model to investigate the potential function of USF in vasopressin promoter activation and development of neuroendocrine SCLC.

We have previously identified a number of enhancers and repressors of the vasopressin promoter [8] and demonstrated that a 199 bp fragment of the human proximal vasopressin promoter is sufficient to retain high-level specific activity in SCLC, but not NSCLC [4]. We identified a neuron-restrictive silencer element (NRSE) motif adjacent to the transcriptional start site [9] that is likely to be involved in this differential regulation via the action of alternative neuron-restrictive silencer factor/RE-1-silencing transcription factor (NRSF/REST) isoforms [10]. In addition, three E-boxes were predicted within this fragment of the promoter [4,11] (Figure 1) and we have shown that the regulatory elements required for high-level expression in SCLC are located between −157 and −131, a region including two of these E-box motifs [12]. E-boxes are known to be important in regulating the tissue-specific expression of several neuropeptides, such as calcitonin-gene-related peptide [13], pro-opiomelanocortin [14] and pre-protachykinin A [15,16]. Members of the basic helix-loop-helix (bHLH) transcription factor family bind E-box motifs and there are a number of candidates for binding at the vasopressin promoter E-boxes in different tissues. These include the product of the proto-oncogene c-*myc*, which has been implicated in neuroendocrine differentiation in SCLC [17] and proneural human achaete-scute homologue 1 (hASH-1), which has been described in SCLC [18,19]. The bHLH-Per-Arnt-Sim (PAS) factors Arnt2 and Sim1 are critical for development of neurons, including the vasopressin secretory neurons of the paraventricular and supraoptic nuclei [20,21]. An E-box of the vasopressin promoter has been found to bind CLOCK/brain and muscle Arnt-like protein 1 (BMAL1; a bHLH-PAS heterodimer) during circadian regulation in the suprachiasmatic nucleus [22].

The USF factors 1 and 2 are encoded by distinct genes and are often regarded as being ubiquitously expressed, although their

Abbreviations used: AVP, arginine vasopressin; bHLH, basic helix-loop-helix; EMSA, electrophoretic mobility-shift assay; hASH-1, human achaete-scute homologue 1; Inr, initiator; NRSE, neuron-restrictive silencer element; NRSF, neuron-restrictive silencer factor; NSCLC, non-small-cell lung cancer; PAS, Per-Arnt-Sim; REST, RE-1-silencing transcription factor; SCLC, small-cell lung cancer; USF, upstream stimulatory factor; USR, USF-specific region.

[1] To whom correspondence should be addressed, at the Physiological Laboratory, University of Liverpool (e-mail j.m.coulson@liv.ac.uk).
[2] Present address: Clinical Oncology, University of Sheffield, Weston Park Hospital, Sheffield S10 2SJ, U.K.

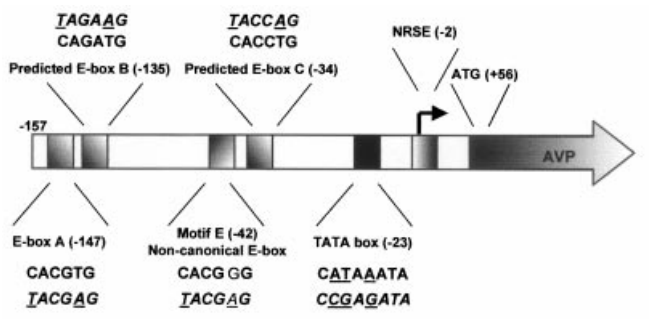

Figure 1 Key motifs within the vasopressin proximal promoter

A map of the proximal vasopressin 5′ gene promoter (not to scale) showing the characterized E-box A, two predicted E-box B and C sequences, the non-canonical E-box (motif E), a characterized NRSE and the TATA box. The 3′ end of all the reporter constructs used in this study is at +42 and the 5′ ends of the proximal (pAVP199) and minimal (pAVP65) vasopressin promoter constructs are at −157 and −23 respectively. The sequences for site-directed mutations described within this paper are shown in italic, with the altered bases underlined. Flanking sequences to E-box C and motif E are shown in Figure 8(A).

expression levels and relative abundance vary [23]. They form hetero- and homo-dimers to bind DNA, although USF-1/USF-2 heterodimers are generally the major USF species [24], with homodimers being less abundant [25]. It is now becoming clear that USF co-operates with other transcription factors to confer tissue-specificity in a number of systems [26–28] and is not ubiquitously active [29,30]. Furthermore, differential splicing generates three human and five rat isoforms of USF-2 that may be important in regulating tissue-specific expression patterns [31,32]. The latter show preferential sequence recognition, and the human USF-2a isoform can function as an activator [26] or play an inhibitory role [33] in the context of different promoters. USF factors have been reported to play key roles in the regulation of a number of promoters; for example USF-1 regulates surfactant A expression in the lung [34]. Interestingly, experiments in null mice show roles for USF-2 in iron homoeostasis [35] and an important role in normal brain function [24]. The latter implies that this 'ubiquitous' factor may in some situations be proneural in a similar way to hASH-1. We predict that such transcription factors may be important in the determination of the SCLC neuroendocrine phenotype. USF factors predominantly function through E-box motifs as transcriptional activators (or repressors). However, their ability to bind other DNA sequences has also been noted; for example there are recent reports of USF binding non-canonical E-boxes [36,37] and these factors can also bind pyrimidine-rich initiator (Inr) sites to act as transcriptional initiators [38].

We have partially characterized E-box A at −147 of the human vasopressin promoter, which bound a bHLH complex in SCLC extracts that was predominantly a heterodimer of USF-1 and USF-2 [12]. Mutational analysis showed that E-box A and an adjacent motif (B), which binds SCLC-specific complexes, together constituted a major enhancer in SCLC [12]. However, USF-1/USF-2 was also bound at E-box A in NSCLC extracts, albeit to a lesser extent than SCLC. In addition, reporter-gene constructs lacking this E-box enhancer retain differential expression between SCLC and NSCLC, and therefore the functional mechanism of USF remains to be elucidated. To determine the role of these complex factors in vasopressin promoter regulation, we report here the effect of overexpressing USF in lung cancer cell lines. This was sufficient to switch on the vasopressin promoter in non-neuroendocrine NSCLC, a cell type where it is otherwise silenced. We have localized the binding sites of USF factors within the vasopressin promoter and, in addition to the −147 E-box, describe a novel non-canonical E-box sequence (CACGGG) at −42 through which activation/initiation by USF is mediated in a TATA-dependent and binding-dependent, but USF-specific-region (USR)-independent, manner.

EXPERIMENTAL

Cell culture

The Lu-165 SCLC line [39] was a gift from Dr T. Terasaki (National Cancer Centre, Tokyo, Japan); the NSCLC line NCI-H460 (large cell), A549 (NSCLC) and NCI-H345 (SCLC) cell lines originated from the ATCC. NSCLC cell lines COR-L23 and MOR/P were gifts from the late Dr P. Twentyman (University of Cambridge, Cambridge, U.K.). All cells were routinely maintained in RPMI/10% bovine calf serum at 5% CO_2 and 37 °C.

Reporter and transcription factor expression constructs

The vasopressin promoter-dependent reporter constructs pAVP65, pAVP173, pAVP185 and pAVP199 are based on the promoterless pBLCAT6 plasmid (pnone) as described previously [4,12]. The locations of predicted E-box and non-canonical E-box motifs within the proximal vasopressin promoter are shown in Figure 1. The reporter constructs mutated within the A and B E-boxes were as described previously [12]. Additional clones corresponding to this region of the vasopressin promoter were generated by PCR amplification and site-directed mutagenesis. The sequence and location of all the primers used are given in Table 1. Forward primers of the same name were used to create the 5′-deleted constructs pAVP119, pAVP105, pAVP94, pAVP94Cmut, pAVP94Emut, pAVP94(EC)mut, pAVP94TATAmut, pAVP80 and pAVP80Cmut. For site-directed mutagenesis of the −34 E-box C, primer pairs in first-round PCR were EBoxC1 with the reverse primer VPP2 and EBoxC2 with the forward primer VPFWD. The products were gel-purified, added in equal quantities as the template in second-round PCR and amplified with VPP2, and either VPFWD or VP186for, to generate pAVP199Cmut and pAVP173Cmut respectively. A similar strategy was used to generate the internal motif E mutation in pAVP199Emut, using the primer pairs VPFWD/VPEMUTrev VPEMUTfor/VPP2, then joining the products from 25 cycles of amplification by annealing and filling in using Platinum *Pfx* DNA polymerase (Invitrogen). All PCR products were ligated into pCR2.1 (Invitrogen) or pDrive (Qiagen) and sequenced before subcloning into pBLCAT6 as described previously for the other constructs [4] or into pGL3-basic (Promega).

The USF expression constructs pN3 (human USF-1) and pN4 (mouse USF-2) and the USF-2 dominant negatives pΔB (capable of dimerization, but lacking a DNA-binding domain) and pXS1 (capable of DNA binding, but lacking the activation domain) were a generous gift from Dr M. Sawadogo (University of Texas, MD Anderson Cancer Centre, Houston, TX, U.S.A.). USF constructs are based on the pSG5 expression vector, which was used as a co-transfection control.

Transient transfection and reporter-gene assays

Single-cell suspensions were prepared either by manual disaggregation (SCLC) or trypsinization (NSCLC); 5×10^6 cells in 400 μl of RPMI/10% bovine calf serum were used for trans-

Table 1 PCR primers used in the generation of reporter gene constructs

Primer sequences are given in a 5′ to 3′ orientation unless otherwise indicated. The position of E-boxes A and C and the non-canonical motif E are shown in bold, the predicted TATA box is in italics and mutated bases are underlined. The *Acc*I cloning site at +42 in VPP2 is in lower case.

Reporter construct	Primer name	Position in AVP gene	Primer sequence
5′ End pAVP199 constructs	VPFWD	−157/−140	CAGACAGGCC**CACGTG**TG
5′ End pAVP173 constructs	VP186for	−131/−110	TGCCTGAATCACTGCTGACGGC
5′ End pAVP119	AVP119	−77/−60	CCACTGGGGAGGGGGTGG
5′ End pAVP105	AVP105	−63/−47	GTGGCGGCCGCGTCTCG
5′ End pAVP94	AVP94	−52/−34	GTCTCGCCTC**CACGGG**AA**C**
5′ End pAVP94Cmut	AVP94Cmut	−52/−26	GTCTCGCCTC**CACGGG**AA**TACCAG**CGG
5′ End pAVP94Emut	AVP94Emut	−52/−31	GTCTCGCCTC**TACGAG**AA**CACC**
5′ End pAVP94 (EC)mut	AVP94(EC) mut	−52/−26	GTCTCGCCTC**TACGAG**AA**TACCAG**CGG
5′ End pAVP94 TATAmut	AVP94TATA mut	−52/−14	GTCTCGCCTC**CACGGG**AA**CACCTG**CGGA*CCGAGATA*GGC
5′ End pAVP80	AVP80	−38/−15	**GG**AA**CACCTG**CGGA*CATAAATA*GG
5′ End pAVP80Cmut	AVP80Cmut	−38/−15	**GG**AA**TACCAG**CGGA*CATAAATA*GG
pAVP199Cmut	EboxC1	−39/−15	5′-**GGG**AA**TACCAG**CGGA*CATAAATA*GG-3′
pAVP173Cmut	EboxC2		3′-**CCC**TT**ATGGTC**GCCT*GTATTTAT*CC-5′
pAVP199Emut	VPEMUTfor	−50/−31	5′-CTCGCCTC**TACGAG**AA**CACC**-3′
	VPEMUTrev		3′-GAGCGGAG**ATGCTC**TT**GTGG**-5′
3′ End: all constructs	VPP2	+36/+57	3′-CCTGGTGCACACAGgtc/gacCC-5′

fection as described previously [4]. Briefly, 10 μg of each reporter plasmid, and 10 μg of transcription factor expression plasmid where appropriate, were delivered at 260 V and 1050 μF. In addition, either 10 μg of pSVβGal or 0.1 μg of pRL-CMV was co-transfected to normalize transfection efficiency. Cell lysates were prepared after a 72 h expression period in 500 μl of lysis buffer. These were assayed for protein content (ESL protein assay; Roche Diagnostics). Chloramphenicol acetyltransferase and β-galactosidase reporter-gene assays used the respective ELISA kits (Roche Diagnostics); pGL3 and pRL-CMV reporter-gene activities were measured using the dual-luciferase assay (Promega).

Preparation of proteins

Whole-cell extracts were prepared from cell lines following three freeze/thaw cycles as described previously [12,40]. Nuclear and cytoplasmic extracts were prepared by selective lysis in Dignam buffer A containing 0.2 % Igepal CA-630, and following centrifugation the supernatant was taken as a cytoplasmic extract and the nuclear pellet lysed in Dignam buffer C. USF expression constructs were translated *in vitro* using the TNT-coupled reticulocyte lysate system (Promega) supplemented with unlabelled methionine.

Electrophoretic mobility-shift assay (EMSA)

EMSA was conducted as described previously [12]. Briefly, double-stranded oligonucleotides were end-labelled with [γ-^{32}P]ATP and reactions carried out in the presence of 500 ng/ml poly(dI-dC) (Sigma) and 10 ng/ml non-specific double-stranded oligonucleotide. Excess of unlabelled competitor double-stranded oligonucleotides or antibodies were pre-incubated with the cell extract or translate for 10 min, then incubated with 1.0 ng of probe at room temperature for 10 min. Samples were analysed on 4 % non-denaturing polyacrylamide (29:1 acrylamide/bisacrylamide) gels in 0.5 × Tris/borate/EDTA buffer, dried and detected by autoradiography. Competitors included a non-specific sequence (NS) [12], E-box A [12] and E-box C (−38/−23; sense strand, 5′-GGGAA**CACCTG**C-GGAC-3′). Other probes/competitors spanned −86/−38, −86/−77, −77/−63, −63/−52, −52/−38, −63/−38, −52/−23 and −45/−30 of the vasopressin promoter, with three bases at either end, beyond that indicated by the numbering. These are illustrated diagrammatically in Figures 6–8 (see below). The −52/−23 Cmut probe contains the 2-bp mutation of E-box C (sense strand, 5′-CGCGTCTCG-CCTC**CACGGG**AA**TACCAG**CGGAC-3′; mutated bases are underlined and the transcription factor-binding motif is shown in bold). The −52/−23 Emut probe contains a similar mutation in the non-canonical E-box motif (sense strand, 5′-CGCG-TCTCGCCTC**TACGAG**AA**CACCTG**CGGAC-3′). Antibodies to USF-1 and USF-2 (N-terminal) were TransCruz supershift antibodies (Insight Biotechnology, Wembley, Middx., U.K.) with 1 μl used per lane in supershift experiments.

Western blotting

Cultured cells were collected at approx. 80 % confluence or transfected cells harvested as whole-cell lysates after a 72 h expression period. Protein extracts from cultured cells or *in vitro* translated proteins were denatured at 100 °C for 5 min in Laemmli buffer (100 mM Tris, pH 6.8, 4 % SDS, 0.2 % Bromophenol Blue, 20 % glycerol and 200 mM dithiothreitol), separated by SDS/PAGE (12 % gels) and electrophoretically transferred to ECL Hybond membranes. Membranes were blocked using 5 % non-fat dried milk in PBS/0.1 % Tween (PBS-T). Primary antibodies were as described for EMSA and used at a dilution of 1 in 5000 in PBS-T. Horseradish peroxidase-conjugated secondary antibodies were used with the ECL system and horseradish peroxidase–streptavidin was used to detect biotinylated markers (Amersham Biosciences). Blots were stripped at 50 °C for 30 min (in 100 mM 2-mercaptoethanol, 2 % SDS and 62.5 mM Tris/HCl, pH 6.7) and re-blocked, before probing with an anti-actin antibody (1:100; Sigma) to control for loading.

RESULTS

USF factors activate the vasopressin promoter in both SCLC and NSCLC

We have characterized endogenous vasopressin expression for a panel of SCLC and NSCLC cell lines [4]. These were used to investigate the contribution of E-boxes and bHLH factors in regulating expression of this neuropeptide gene. USF-1 and

USF-2 expression constructs were transiently co-transfected with vasopressin promoter constructs into two cell lines: Lu-165 (SCLC), which has a very high level of vasopressin expression [39], and NCI-H460 (NSCLC), in which the vasopressin transcript was not detected [4]. Reporter-gene expression dependent on the pAVP199 vasopressin promoter is shown in Figure 2(A). Whereas USF-1 had no effect on pAVP199 activity in the SCLC cell line, USF-2 significantly enhanced this by over 2-fold ($P < 0.005$), providing functional confirmation that this factor binds an enhancer domain in the proximal vasopressin promoter.

The most striking observation, however, was that both USF-1 and USF-2 could activate the vasopressin promoter in the NSCLC line (NCI-H460), where it is normally silenced (Figure 2A). This effect was most marked with USF-2, which conferred activity equivalent to the high basal activity of the Lu-165 SCLC cell line. The activation was not due to cross-talk between promoters, as neither of the two USF-2 mutants (pΔB and pXS1), which are binding- and activation-deficient respectively, nor the pSG5 empty expression vector, caused activation of the pAVP199 vasopressin promoter in NCI-H460 (Figure 2B). Although NSCLC comprises a heterogeneous group of lung tumours, we have previously demonstrated a lack of endogenous vasopressin production in a panel of NSCLC lines and shown that vasopressin promoter-driven reporter constructs are negligibly expressed [4]. Activation of the vasopressin promoter by USF factors was tested in three other NSCLC cell lines (Figure 2C). In each line there was little or no basal pAVP199 reporter-gene expression, but the construct was substantially activated by co-transfected USF-2. There was no clear additive or synergistic effect on co-transfection of USF-1 with USF-2, implying that activation can be mediated by a homodimer.

Expression of USF in lung tumour cell lines

The ability of transfected USF to activate the vasopressin promoter in NSCLC implies that endogenous levels, or isoforms of the factor, may vary between SCLC and NSCLC. USF-1 and USF-2 expression was investigated by Western blotting; representative data for Lu-165 (SCLC) and NCI-H460 (NSCLC) are shown in Figure 3(A). No USF-1 expression was detected in the cell lines at this exposure, although some USF-1 was evident in the SCLC nuclear extract at long exposures (results not shown). However, the antibody clearly recognized *in vitro*-translated USF-1 (Figure 3A), implying that this factor was of very low abundance in lung cancer cells. This must, however, be sufficient for formation of the heterodimer that binds E-box A [12], but may be a limiting factor for this enhancer activity.

USF-2 was more highly expressed and was localized to the nucleus in SCLC. The full-length 44 kDa form was predominant, with little evidence of shorter isoform expression (although the antibody against the N-terminal would recognize the two major forms). Western blots were re-probed with an anti-actin antibody, and actin was evident in both cytoplasmic and nuclear fractions. Densitometry was used to calculate the ratio of USF-2 to actin for each sample, as shown in Figure 3(B), to standardize the cell extracts. In contrast with SCLC, very little USF-2 was detected in the NSCLC nuclear fraction (7% of that in SCLC). Although USF-2 was detected in whole-cell lysates of NSCLC, this represented a very low level when standardized against actin. In contrast, the relative USF-2 level was increased by 6.5-fold in whole-cell extract prepared from the same cells transfected with USF-2. Therefore USF protein expression in NSCLC is very low compared with that in SCLC and can be substantially increased by transfection of the USF-2 expression plasmid pN4, correlating to transcriptional activity of the vasopressin promoter-driven reporter-gene constructs.

USF factors demonstrate enhancer activity through E-box A

We have generated a number of vasopressin promoter 5′-deletion or mutated constructs, which remain inactive in NSCLC cell lines when transfected alone or with the pSG5 empty expression vector (results not shown). USF expression constructs were co-transfected into NCI-H460 (NSCLC) in combination with this initial series of reporter-gene constructs (Figure 4A) to determine whether the previously predicted E-box motifs were mediating activation. These data were standardized internally against pAVP199 co-transfected with either USF-1 (Figure 4B) or USF-2 (Figure 4C), but co-transfection with USF-1 reproducibly resulted in around 50% of the activity for USF-2, as shown in Figure 2(A).

Both USF-1 (Figure 4B) and USF-2 (Figure 4C) were clearly capable of conferring transcriptional activity in NSCLC for most of the promoter fragments, but there may be several distinct roles for USF in vasopressin promoter regulation. On comparison of pAVP199 and pAVP199(AB)mut, it appears that the AB enhancer we have described previously in SCLC [12] can also function in NSCLC when USF factors are overexpressed, giving an approx. 2-fold enhancement in reporter-gene expression (Figure 4). Binding studies in SCLC indicated this would be mediated by a USF-1/USF-2 heterodimer [12]. We now show that this can, at least *in vitro*, be substituted by USF-1 or USF-2 homodimers in NSCLC. Mutation of any E-box within the pAVP199 fragment reduced USF-2 activation (Figure 4C). We have previously shown that the predicted E-box B weakly binds specific complexes in SCLC, which may modulate USF enhancer function via E-box A [12]. Comparison of pAVP199 with pAVP199Bmut shows that the B motif affected enhancer activity in NSCLC for an USF-2, but not an USF-1 homodimer, with a significant difference ($P < 0.005$) between the two factors (Figure 4C). These data may point to USF-2 as the important heterodimer partner that determines expression levels through E-box A in SCLC. Clearly, USF factors can differentially regulate through this complex E-box enhancer when overexpressed in NSCLC.

USF activation requires −52 to +42 of the vasopressin promoter

Although USF factors could bind the AB enhancer motif, this was not the major or sole determinant of promoter activation by USF in NSCLC, as the promoter was still switched on where this E-box was deleted (pAVP173) or mutated [pAVP199(AB)mut; Figure 4]. Mutation of E-box C in the context of pAVP199Cmut or pAVP173Cmut (in which all the previously predicted E-boxes have been deleted or mutated) reduced activity to a level comparable with pAVP199Amut (Figure 4), implying some synergy or interaction of E-box C with other motifs. However, USF factors failed to activate both the pAVP65 minimal promoter fragment or pAVP80. In contrast, for SCLC both pAVP65 and pAVP80 are active and mutation of E-box C had no effect (results not shown). As pAVP80 retains E-box C, this demonstrates the requirement of additional sequence for USF activation in NSCLC (Figure 4).

It therefore appeared that the vasopressin promoter activation was not solely dependent on a canonical E-box, but that there was a requirement for the promoter region from −86 to −38, which defines the 5′ ends of the constructs pAVP128 and pAVP80 (Figure 4). We generated three further deletion constructs to

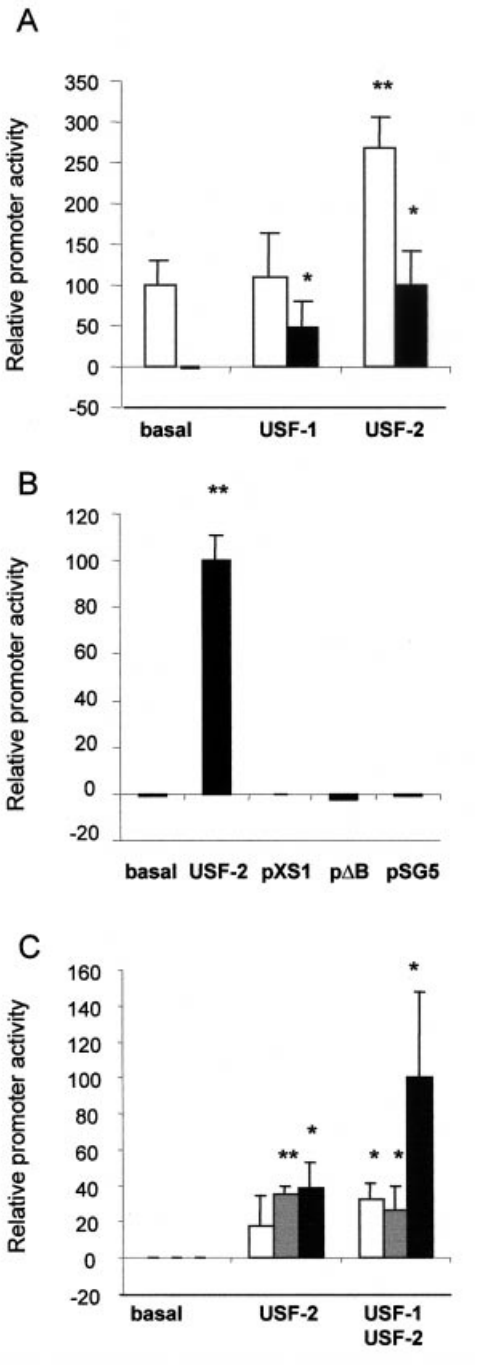

Figure 2 The effect of co-transfected USF expression constructs on vasopressin promoter activity in lung cancer cell lines

In each case the mean ± S.D. ($n = 3$) is shown and statistical significance determined by Student's t test is indicated by ** ($P < 0.005$) and * ($P < 0.05$). (**A**) Reporter-gene expression on transfection with pAVP199 alone (basal) and for co-transfection of pAVP199 and transcription factor expression constructs for USF-1 or USF-2 in Lu-165 (SCLC; white bars) and NCI-H460 (NSCLC; black bars) cell lines. Reporter-gene activity is expressed as a percentage of the basal pAVP199 promoter activity in SCLC and significance indicated relative to the basal pAVP199 level for each cell line. (**B**) The effect of USF-2 dominant negatives (pXS1 and pΔB) and empty expression vector (pSG5) on the pAVP199 vasopressin promoter in NCI-H460 (NSCLC) cells. Reporter-gene activity is expressed as a percentage of the USF-2-stimulated level and significance relative to basal is indicated. (**C**) The effect of USF on pAVP199 vasopressin promoter activation in other NSCLC cell lines: COR-L23 (large cell; white bars), MOR/P (adenocarcinoma; shaded bars) and A549 (adenocarcinoma; black bars). Unstimulated basal promoter activity is shown compared with co-transfection of USF-2 alone, or of USF-1 and USF-2 together. Data are expressed as a percentage of the activity for A549 with both USF expression constructs, and significance is indicated relative to the basal level for each cell line.

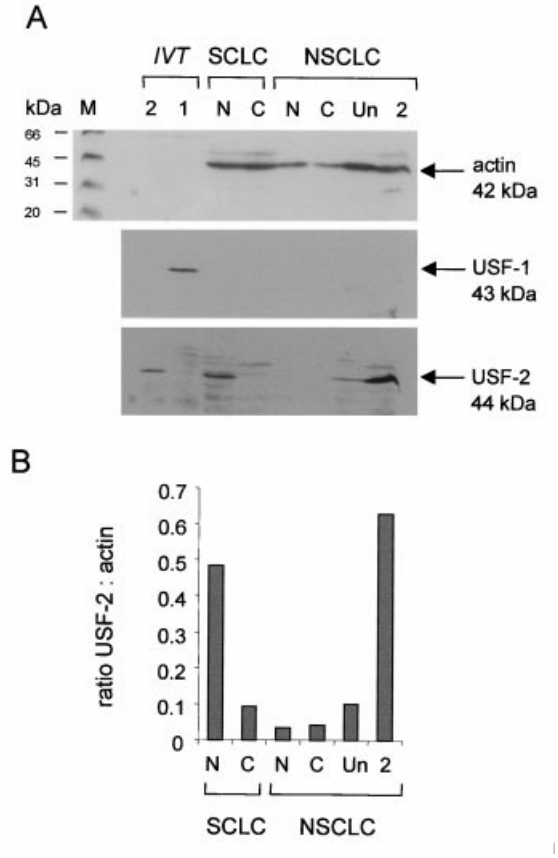

Figure 3 Expression of USF in lung tumour cell lines

(**A**) A Western blot probed with antibodies for actin (top), USF-1 (middle) and USF-2 (bottom). Samples are: *IVT*, *in vitro*-translated USF-2 (lane 2) and USF-1 (lane 1); SCLC, nuclear (lane N) and cytoplasmic (lane C) extracts from Lu-165; NSCLC, nuclear or cytoplasmic extracts and whole-cell extracts from untransfected control (lane Un) or USF-2-expression-construct-transfected (lane 2) NCI-H460 cells. (**B**) Histogram showing the ratio of USF2 to actin determined from the above Western blot. ECL autoradiograms were imaged as TIF files, imported into Scion Image (Scion Corp.) and the pixels associated with each band determined.

dissect this intervening region. The USF-2 binding deficient mutant pΔB did not activate reporter-gene expression from any of these constructs (Figure 5), showing that USF-2 binding is required for function. In a series of experiments these were co-transfected with USF-2, USF-1 or USF-1 and USF-2, the latter having potential for heterodimer formation. Comparable activation was seen for both heterodimers and homodimers. We observed a pattern of reduced activity at each consecutive deletion from the 5′ end of the promoter, but activation was only lost on deletion from pAVP94 (-52 to $+42$) to pAVP80 (-38 to $+42$; Figure 5). Thus activation requires USF binding to a site between -52 and -38, functionally excluding E-box C.

USF factors bind E-box A and a non-canonical E-box motif, but not E-box C

We have previously shown by EMSA that E-box A (-147) differentially binds USF heterodimers in SCLC and NSCLC. We used *in vitro*-translated USF to confirm whether these factors bind as homodimers to either E-box A or E-box C ($-38/-23$, C on Figure 6A). Whereas USF-2 formed a complex on the

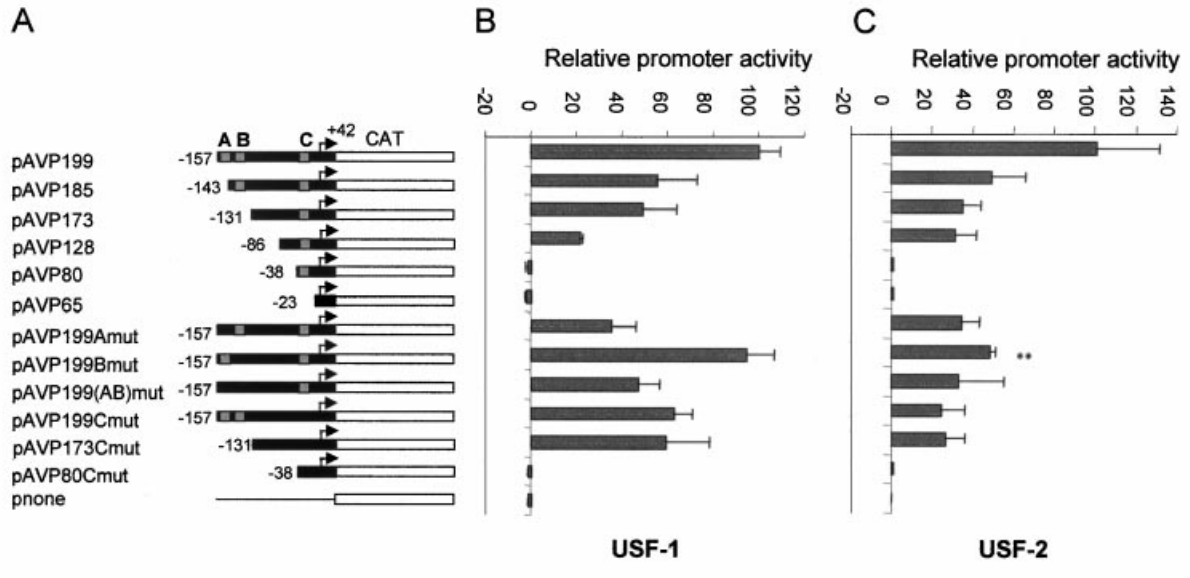

Figure 4 The contribution of previously identified E-box motifs to USF-mediated activation of the vasopressin promoter in NSCLC

(**A**) The chloramphenicol acetyltransferase reporter constructs used include 5′ deletions and mutations based on the pAVP199 vasopressin promoter; these are illustrated diagrammatically showing the intact E-boxes (A–C) in each case. (**B** and **C**) Vasopressin promoter activity expressed as a percentage of that for pAVP199, on co-transfection with USF-1 and USF-2 expression constructs. In each case the mean ± S.D. ($n = 3$) is shown. Statistical difference between the relative results for a reporter construct co-transfected with either USF-1 or USF-2 was determined by Student's t test and is indicated in (**C**) by ** ($P < 0.005$).

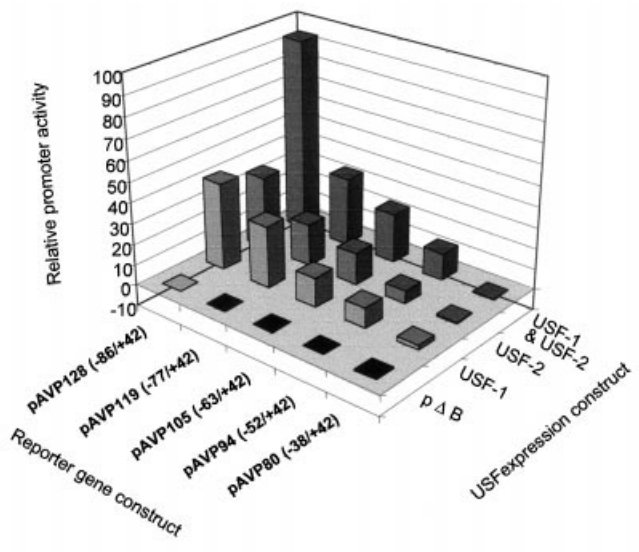

Figure 5 Identification of the promoter region required for USF activation in NSCLC

Three additional deletion constructs were generated to represent the intervening region between pAVP128 (−86 to +42) and pAVP80 (−38 to +42); these were co-transfected into NCI-H460 cells (NSCLC) with USF factors (USF-1, USF-2, USF-1 and USF-2 together and pΔB) to locate the activator region. Reporter-gene activity is expressed as a percentage of that for pAVP199 co-transfected with USF-2 ($n = 3$; S.D. $\leqslant 15\%$).

former, no specific binding was observed on the latter, as determined by comparison with the non-specific competitor sequence (NS). It appears that E-box C in isolation does not bind full-length USF-2, concurring with the functional data that pAVP80, spanning (−38 to +42) and containing E-box C, was not activated by USF. Instead, this E-box was seen to bind an unknown complex in NSCLC (results not shown). Therefore the region from −52 to −38, which we had shown to be functionally important for USF initiation of the vasopressin promoter, must mediate the effect via an alternative site.

EMSA was used to investigate direct binding of USF within this region, by competition for *in vitro*-translated USF-2 bound to the E-box A probe. The vasopressin promoter from −86 to −23 is highlighted in Figure 6(C), with the series of double-stranded competitor oligonucleotides, representing the 5′ ends of the different reporter constructs, shown below. The E-box C probe −38/−23 failed to compete for bound USF-2; indeed, the only competition was seen with the −52/−23 probe, although this was incomplete at a 100-fold excess over labelled E-box A probe (Figure 6B). Identical results were seen using an extract from NSCLC cells transfected with the USF-2 expression construct (Figure 6B). In both protein extracts, the bound complex was confirmed as a USF-2 homodimer by antibody supershift. The −52/−38 probe, like −38/−23, does not compete for USF-2, implying that the binding site overlaps these two probes.

In the converse experiment to that described above, *in vitro*-translated USF-2, or extracts from NSCLC cells transfected with USF-2, showed formation of a specific complex on the labelled −52/−23 probe (Figure 7B). This complex was competed by E-box A but not −38/−23 (E-box C) at 100-fold excess, displayed the same mobility as USF-2 and was supershifted by an USF-2 antibody. In contrast, no complex was bound to −52/−23 in control NSCLC extract (Figure 7B), implying that insufficient endogenous USF-2 is present in NSCLC to bind at this site, unless USF-2 is overexpressed from the transfected pN4 construct. Interestingly, USF-2 was bound by the −52/−23 probe

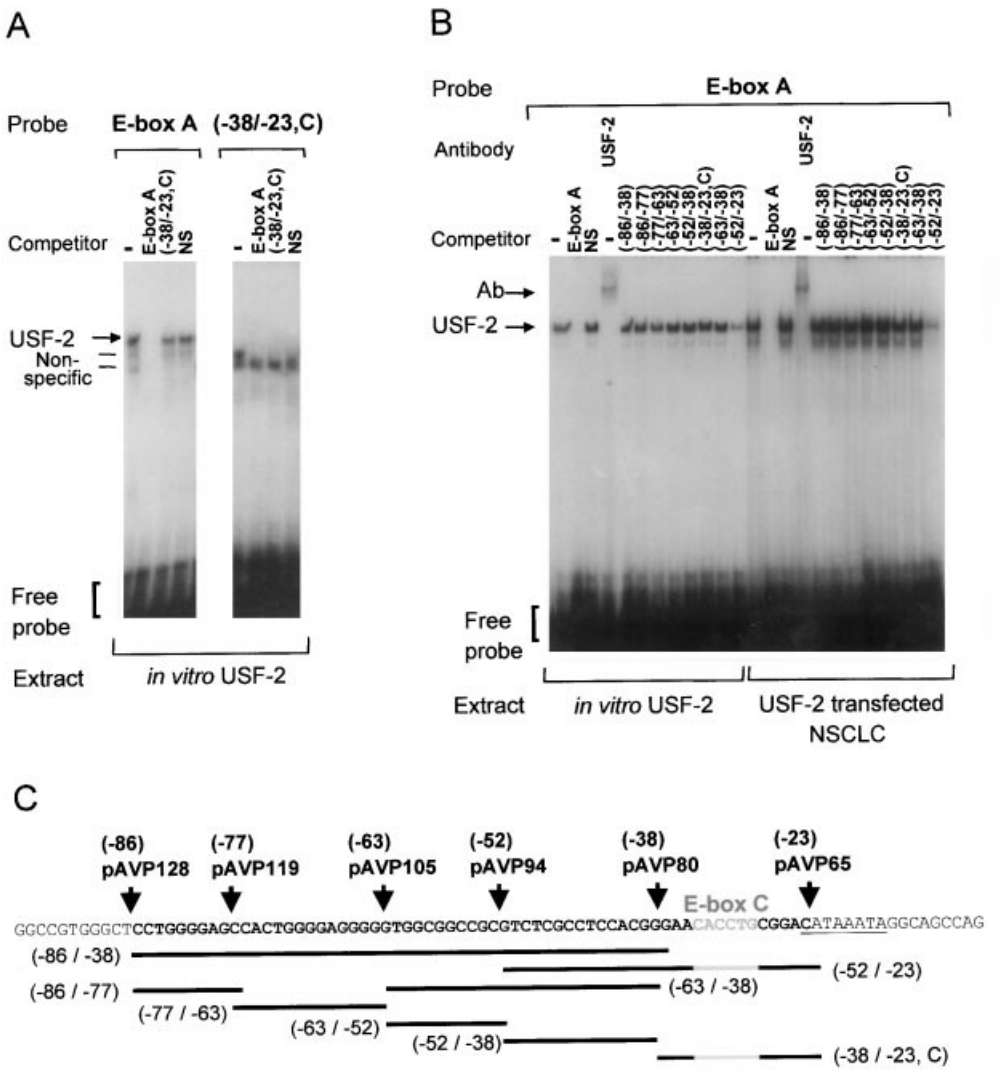

Figure 6 EMSA of USF-2 binding to the vasopressin promoter

(**A**) Binding of *in vitro*-translated USF-2 to E-boxes A and C (−38/−23, C); specific binding (indicated by an arrow) is seen to probe A, some non-specific binding (indicated by lines) is seen for both probes. Competition is shown with a 100-fold excess of unlabelled oligonucleotide as indicated. (**B**) Binding of USF-2 from *in vitro*-translate or USF-2-transfected NCI-H460 cells (NSCLC) to the E-box A probe is competed with a 100-fold excess of unlabelled oligonucleotides. These correspond to fragments of vasopressin promoter shown below; only −52/−23 partially competes for bound USF-2. (**C**) The sequence of the vasopressin promoter from −97 to −7, highlighting the region of interest for USF-2 activation in bold and showing E-box C at −34, the TATA box (underlined) and the EMSA probes.

in an extract from the Lu-165 SCLC cell line (Figure 7B), where both endogenous USF-2 (Figure 3) and the vasopressin promoter [4] are known to be highly expressed. In these SCLC cells two additional specific complexes were also seen, which were competed by an excess of unlabelled −52/−23 probe, but not by the E-box A probe, and were not supershifted by the USF-2 antibody. Therefore in SCLC, where vasopressin is expressed, both USF-2 and additional factors bind within this promoter region. In both USF-2-transfected NSCLC and untransfected SCLC, there was only weak competition for the USF-2 complex bound to −52/−23 by excess −38/−23 (E-box C). The −52/−23 probe also includes the E-box C motif, and an identical probe mutated within this E-box (−52/−23 Cmut; Figure 7A) retained the ability to compete for USF-2 (Figure 7B). Thus direct USF interaction with the vasopressin promoter is largely independent of the −34 E-box C, but correlates to the vasopressin expression status of the cells.

USF-2 binds and activates via a vasopressin promoter CACGGG motif (−42 to −37)

A non-canonical E-box, termed motif E, was identified within the key region of the vasopressin promoter required for USF functionality and binding in this model (Figure 8A). The sequence CACGGG differs by only one base from the preferred USF-binding site of CACGTG. Motif E spans the ends of the −38/−23 and −52/−38 probes, neither of which efficiently competed for USF-2 binding (Figures 6B and 8B). However, the −45/−30 probe contains motif E (Figure 8A) and competed with a similar efficiency to −52/−23 (Figure 8B). Mutation of

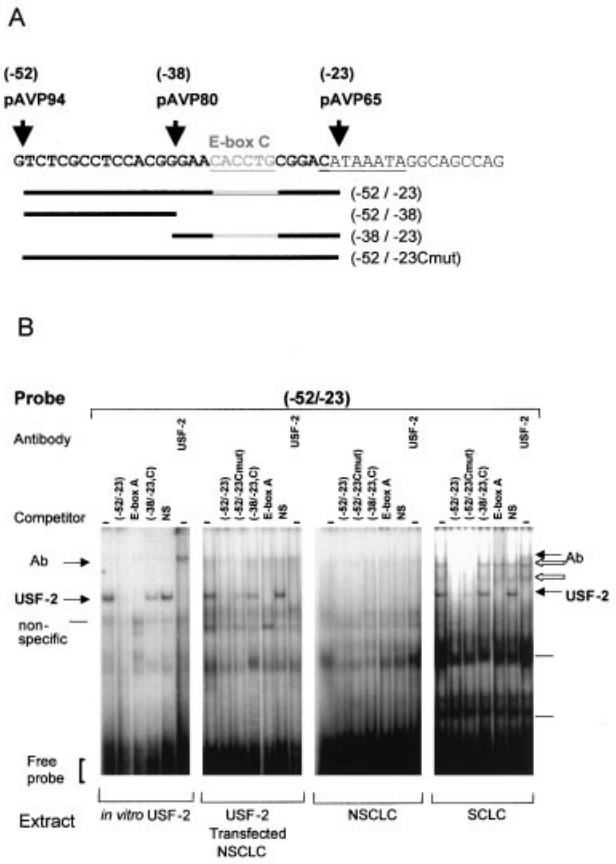

Figure 7 USF-2 binds to the −52/−23 probe, but not via E-box C

(**A**) The vasopressin promoter from −52 to −7, showing EMSA probes below with the −34 E-box C highlighted where present. (**B**) Binding of USF-2 to the −52/−23 probe for *in vitro*-translated USF-2, USF-2-transfected NCI-H460 cells (NSCLC), control untransfected NCI-H460 cells or Lu-165 cells (SCLC). Competition is shown with a 100-fold excess of unlabelled probes as indicated. Arrows indicate specific complexes (USF-2) and antibody supershifts (Ab), with non-specific complexes shown by single arrows and SCLC-specific complexes by white arrows. The specific complex is recognized by a USF-2 antibody and competed efficiently by the −52/−23 and E-box A probes, but not the E-box C probe (−38/−23, C). USF-2 binding is not detected in untransfected NSCLC, while USF-2 and two additional complexes form on the probe with the SCLC extract.

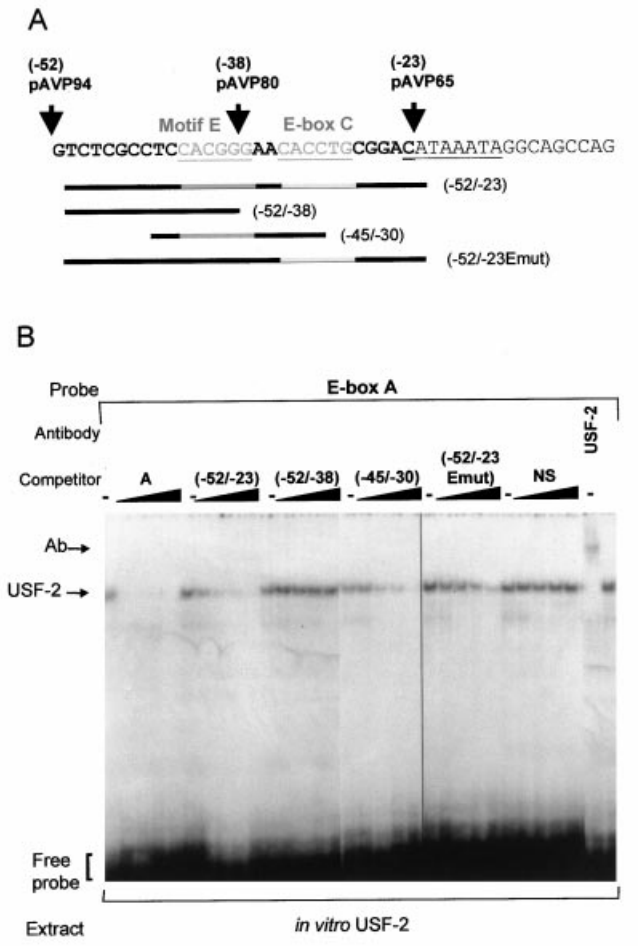

Figure 8 Identification of a CACGGG binding motif for USF-2

(**A**) The sequence of the vasopressin promoter from −52 to −7, showing E-box C and the non-canonical E-box (motif E) within the EMSA probes used. (**B**) EMSA showing competition for *in vitro*-translated USF-2 bound to the E-box A probe with titration of 0, 10, 50, 100 and 500-fold excess of competitor. Oligonucleotides −52/−23 and −45/−30 compete, whereas −52/−38 and −52/−23 (Emut) compete less efficiently, confirming motif E (−42) as the major binding site for USF-2 in this region.

motif E within the context of −52/−23 severely impaired the ability of this probe to compete for bound USF-2, with competition still incomplete at 500-fold excess (Figure 8B).

Titration of competition for USF-2 bound to the E-box A probe with homologous oligonucleotide or the −52/−23 probe illustrated that, while the latter binds specifically, it does so with a substantially lower affinity than the canonical E-box A (Figure 8B). Complete competition is seen with a $\leqslant$ 10-fold excess of E-box A probe, but a 100-fold excess of −52/−23 is required. We have shown that USF-2 is expressed at very low levels in the nuclear fraction of NSCLC cells (Figure 3), but that this can bind to a limited extent via the E-box A enhancer [12]. However, endogenous USF-2 in NSCLC clearly does not bind to motif E (Figure 7B), and the low affinity of USF-2 for this non-canonical E-box may in part account for the normally inactive state of the vasopressin promoter in NSCLC. These data point to motif E as the key site for USF-2 binding.

Functional confirmation was obtained by introducing mutations of motif E, E-box C and the non-canonical TATA box into pAVP94, the minimal reporter-gene construct that supports USF activation. The TATA mutation is based on the previously published mutation of a canonical TATA box used in conjunction with USF-2 initiation studies [41]. The mutated bases are common between the canonical TATA (TATAAATA mutated to TCGAGATA) and the AVP non-canonical TATA (CATAAATA mutated to CCGAGATA). In concordance with the previous binding and functional data, mutation of E-box C does not affect USF-2 activation. However, mutation of either motif E or the TATA box abolished activation, reducing levels to that of the minimal pAVP65 construct that does not support USF-2 activation in this model (Figure 9A). Thus it appears that the vasopressin non-canonical TATA box is functional and important for initiation in this model, but that motif E is also essential for activation of the TATA complex in NSCLC. Motif E was then mutated in the context of the pAVP199 promoter construct to determine whether this non-canonical E-box operates in isolation from, or conjunction with, E-box A (Figure 9B). In common with the other vasopressin promoter constructs,

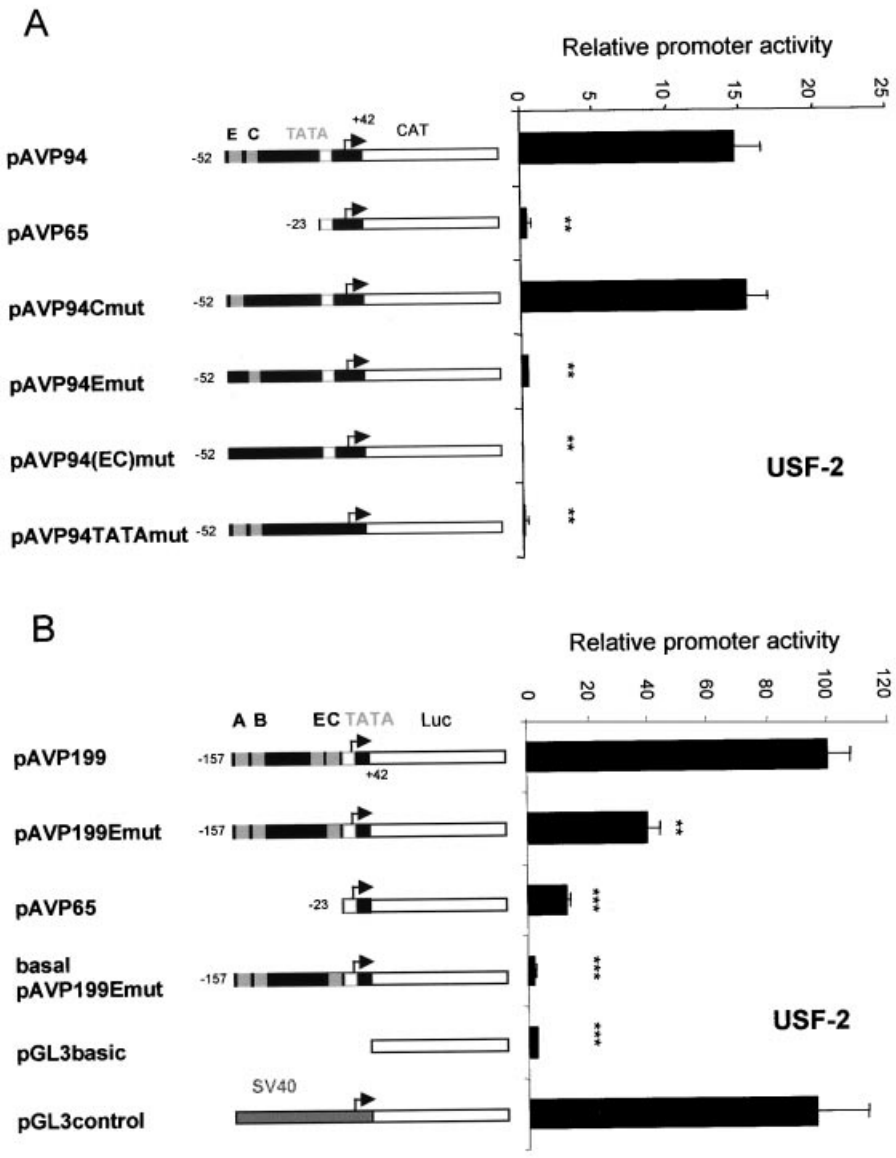

Figure 9 USF-2 activates via the CACGGG motif E in pAVP94 and pAVP199 constructs

(**A**) Chloramphenicol acetyltransferase (CAT) reporter-gene constructs dependent on the pAVP94 promoter were mutated as described in the Experimental section and are illustrated diagrammatically. These were co-transfected into NCI-H460 (NSCLC) cells with the USF-2 expression construct. Data are represented as a percentage of the activity for pAVP199 co-transfected with USF-2, the mean ± S.D. ($n = 3$) is shown; Student's t test is indicated as ** ($P < 0.005$). Mutation of E-box C does not affect USF activation, whereas mutation of either motif E or the non-canonical TATA box abolishes initiation. (**B**) A luciferase (Luc) reporter-gene construct dependent on the pAVP199 promoter was mutated at motif E. Co-transfections were carried out in NCI-H460 cells (NSCLC) with the USF-2 expression construct (apart from basal pAVP199Emut). Data are represented as a percentage of the activity for pAVP199 co-transfected with USF-2, the mean ± S.D. ($n = 3$) is shown, and Student's t test is indicated as ** ($P < 0.005$) and *** ($P < 0.001$). Motif E clearly contributes to USF-2 activation of p199AVP.

basal transcription from pAVP199Emut in the NSCLC line is minimal. However, when pAVP199Emut was co-transfected with USF-2, it was activated to only approx. 40 % of pAVP199. Motif E would therefore appear to play a key role in determining whether the vasopressin gene is expressed and to synergize with E-box A. In contrast with these data, preliminary analysis of reporter-gene expression in SCLC cells indicates that mutation of motif E had no effect on the activity of the pAVP199 proximal promoter (results not shown). This is not unexpected, as the transcription factor complement of SCLC and NSCLC may differ substantially; indeed, EMSA data indicated several other factors in SCLC extracts with binding sites in this region (Figure 7B).

Proposed mechanism for USF initiation

Reporter plasmids were co-transfected with equivalent amounts of USF-2 (pN4) and the expression constructs for activation-deficient (pXS1) or binding-deficient (pΔB) USF-2. Both dominant negative heterodimers clearly had a significant effect on the pAVP199 construct, reducing activation to around 20–40 % of that seen with USF-2 alone (Figure 10A). The pAVP94 promoter

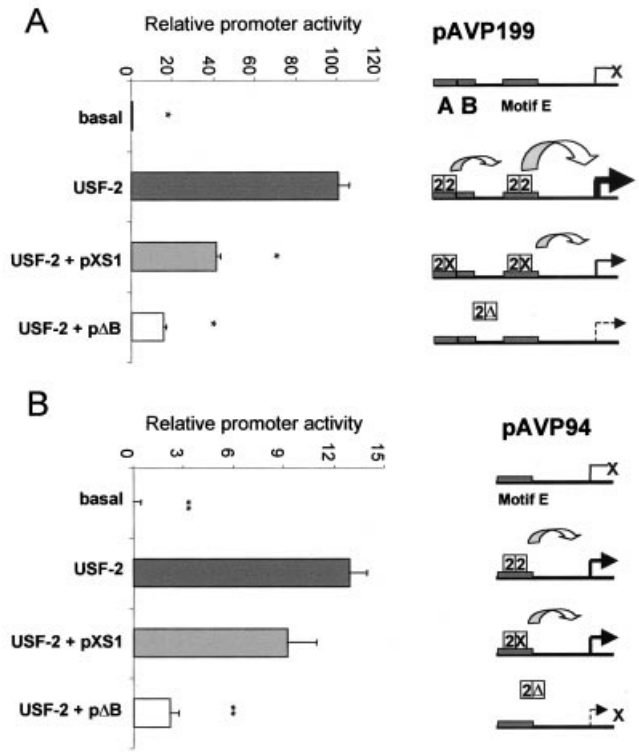

Figure 10 USF activation via motif E requires the binding domain of USF-2

The effect of USF dominant negatives on USF-2 activation of vasopressin promoter-dependent reporter-gene expression in NCI-H460 cells (NSCLC) is shown for (**A**) pAVP199 (−157 to +42) and (**B**) pAVP94 (−52 to +42). Different dominant negative expression constructs were co-transfected at equimolar ratios with pN4, which expresses USF-2. In each case the reporter construct plus the empty expression vector pSG5 is shown as basal. Means ± S.D. are shown for promoter activity relative to pAVP199 co-transfected with USF-2 ($n = 3$). The effect of co-transfection with the activation-defective USF-2 mutant pXS1, or the binding-defective USF-2 mutant pΔB, are compared (Student's t test is indicated as *, $P < 0.05$, and **, $P < 0.005$). Proposed models for USF activation of the pAVP199 and pAVP94 promoter fragments via E-box A (−147) and motif E (−34) are shown to the right. Heterodimers are shown as binding on co-transfection, although homodimers of USF-2 may also form, resulting in some leakiness of inhibition. Key: 2, USF-2; X, pXS1; Δ, pΔB.

(−52 to +42) co-transfected with USF-2 exhibited around 10% of the reporter-gene activity of pAVP199 and, in contrast to pAVP199, this was not significantly reduced by co-transfection with pXS1 (Figure 10B). However, the binding function of USF-2 was clearly essential for activation of the pAVP94 construct, as reporter-gene expression was lost almost completely on co-transfection with pΔB (Figure 10B).

These data are interpreted in the adjacent schematics where, for pAVP199, USF-2 binds through both E-box A and motif E, but for pAVP94 only the latter motif is present. While the binding-deficient mutant can disrupt USF-2 function via both motifs, only the E-box A enhancer is affected by USF-2 heterodimerization with the activation-deficient mutant. Therefore, the USR activation domain of USF-2 [41] does not appear to be required for activation, although pXS1 alone does not activate the vasopressin promoter (Figure 2B). However, binding at motif E was required, and other USF domains, or interaction with other factors, may be involved in its function through the non-canonical E-box. Our data show that USF-2 preferentially binds E-box A over motif E *in vitro* and that the low level of endogenous USF in NSCLC can complex at the E-box A enhancer, but not motif E. Thus, overexpression of USF appears to lead to binding at the putative initiation motif E, overriding the normal repression of the vasopressin promoter reporter constructs in NSCLC.

DISCUSSION

In this study we have elucidated a complex role for USF in regulation of the vasopressin promoter in a lung cancer model, via both an E-box and a non-canonical E-box motif. It is likely that these and other motifs are involved in physiological vasopressin regulation through interaction with USF or alternative bHLH or bHLH-PAS factors. Tissue-specific or pathological roles of USF are increasingly being defined [42,43] and the original concept of these as ubiquitous transcriptional regulators is changing. USF-2 is expressed at relatively low levels in normal adult lung [44], and our previous EMSA data indicated that USF factors may function differentially in lung cancer [12]. This could reflect a variation in either the USF expression level or USF isoforms between SCLC and NSCLC. We found quantitative differences in the major 44 kDa isoform, with abundant USF-2 in SCLC and much lower levels in NSCLC. Interestingly, this expression differential between SCLC and NSCLC adenocarcinoma is borne out in larger studies of lung cancer cell lines and clinical tumour material (M. Ocejo-Garcia, J. M. Coulson, I. Somoro and P. J. Woll, unpublished work). USF functionality may of course be regulated at other levels, such as phosphorylation that can modulate DNA binding [45] and by interactions dependent on the complement of co-activators [30], other bHLH factors, or Id proteins, within cells.

USF has been reported to be anti-proliferative, sometimes with an antagonizing function to c-Myc [46], for example activating transcription of the mitogen-activated protein kinase regulator Mkp-1 [47] and the adenomatous polyposis coli (APC) tumour-suppressor gene [48]. Thus the low expression of USF-2 seen in NSCLC is perhaps more typical of cancer cells. Interestingly, a loss of transcriptional activity, rather than a change in protein levels, has also been reported for USF in epithelial-derived breast cancer cell lines [29] and in other tissue-specific expression patterns [28]. Therefore, the low level of USF proteins we report here in NSCLC, and the weak complex we have previously shown at E-box A in NSCLC extracts [12], may not represent transcriptionally active USF. In neuroendocrine SCLC, which display less epithelial-like characteristics, we have clearly demonstrated that USF-2 is highly expressed and transcriptionally active. This is probably only one of many transcriptional differences between SCLC and NSCLC.

Overexpression of USF in NSCLC activated the vasopressin promoter, and a component of this enhancer activity was shown by site-directed mutagenesis and EMSA to be mediated via E-box A, as previously proposed in SCLC [12]. However, other motifs were also involved, and we had postulated that E-box C might be important in modulating vasopressin promoter activity, due to its location at −34 relative to the transcription initiation site. In fact, E-box C was insufficient for USF-2 activation in NSCLC and did not directly bind USF, but may instead bind a repressor. There are several other mechanisms by which USF could exert its function; for example, indirectly, as described in transcription of L-type pyruvate kinase [49]. Instead, we found that USF directly bound the vasopressin promoter via a previously uncharacterized non-canonical E-box (CACGGG, motif E) located at −42. Several functional non-canonical E-boxes have recently been described that contribute to stress-induced regulation of target genes, for example activation via the heptameric CAGCCTG motif in the lama 3 promoter [37] and

repression via the methylation-sensitive AACGTG motif in the fibroblast growth factor-binding protein promoter [36]. It is likely that in the vasopressin promoter, the flanking region of motif E could be disrupted by mutations generated in the adjacent E-box C, which may in part explain the disparity between functional data for p199Cmut and evidence that USF-2 does not bind E-box C.

USF is related to the basal transcription factor TFII-I and can bind pyrimidine-rich Inr sites in addition to E-boxes [38]. Inr motifs can either initiate transcription independently or increase activation via a TATA box [50], and USF has been reported to initiate certain promoters via such an Inr [51], which again can be dependent on a TATA box [41]. USF initiation has also been described through an E-box motif in a gene lacking both a TATA box and Inr, suggesting that E-boxes can be core elements that direct basal transcription [52]. Recently, a role for USF was reported in the dipeptidyl peptidase IV promoter, where USF acts through both an E-box enhancer and a non-E-box initiator motif [53]. A 'TATA' box (CATAAATA) was originally predicted within the human vasopressin promoter [54], but has not been unequivocally shown to act as a non-canonical TATA motif. However, it was recently shown that this 'CATA' box, although weaker than a canonical 'TATA' box in HEK-293 cells, interacts more efficiently with an upstream enhancer and differentially binds complexes in supraoptic nucleus extracts [55]. This 'CATA' box is present at the 5′ end of the minimal pAVP65 (-23 to $+42$) reporter construct. pAVP65 shows basal activity in SCLC, but is inactive in NSCLC even when USF-2 is overexpressed. However, based on our mutational analyses in NSCLC, USF activated the pAVP94 (-52 to $+42$) reporter constructs in a CATA-dependent manner via the non-canonical motif E. Therefore, motif E (-42) clearly plays a major role along with E-box A (-147) in USF activation of the vasopressin promoter, possibly modulating initiation via the CATA box. This arrangement of a consensus E-box and other non-canonical USF binding motifs is comparable with the dipeptidyl peptidase IV promoter architecture.

The activation mechanism is likely to not only involve these co-operative interactions between E-boxes and the basal transcription machinery, but also interactions with other transcription factors. There are a number of potential transcription factor-binding sites in this proximal promoter region, including a GC-rich region described previously in the rat vasopressin promoter spanning -94 to -60 [11] and implicated as a repressor of the bovine vasopressin promoter [55]. Other predicted sites include LBP-1 (also known as upstream binding protein 1) and MAZ (Myc-associated zinc finger protein), both mapping to the region between the 5′ ends of pAVP119 and pAVP105 (-77 to -63), and YY1. Although this region was not essential for USF activation in NSCLC, it can potentially bind other complexes in SCLC, and EMSA data from the $-52/-23$ probe show that additional DNA-binding complexes are present in SCLC cells where the endogenous gene is transcribed.

The vasopressin promoter architecture also suggests that a NRSE-like repressor element, which we have described at the transcriptional start site, may regulate initiation mechanisms [9]. We have previously shown that the pAVP65 minimal promoter fragment (-23 to $+42$) is sufficient to retain promoter activity in SCLC [4] and this may be disregulated by an isoform of NRSF/REST in this tumour type [9,10]. In USF-2-transfected NSCLC the NRSE motif may maintain the repressed state of the pAVP65 (-23 to $+42$) and pAVP80 (-38 to $+42$) constructs, but when motif E is present USF-2 can overcome this repression. We have generated preliminary data, by reverse-transcriptase PCR on NSCLC cells transiently transfected with the USF-2 expression construct, which indicated that vasopressin mRNA was not present at detectable levels 3 days post-transfection (results not shown). Therefore transcription from the endogenous vasopressin gene does not appear to be up-regulated in USF-2-transfected NSCLC in the same way as the reporter-gene constructs and it is likely that more complex interactions of factors are required to fully activate the chromatin-associated structural gene. NRSF is known to interact with histone deacetylases [56,57] and it has been shown that USF can overcome transcriptional repression by displacing nucleosomes located at a promoter region [58]. It may be that such chromatin-modifying functions play a part in determining the balance of endogenous vasopressin promoter expression in SCLC and that these functions are not fully supplied by USF-2 overexpression alone in NSCLC.

Both heterodimers and homodimers of USF factors activated the vasopressin promoter. This contrasts with findings in other systems where for example overexpression of USF heterodimers activated, but homodimers inhibited, transcription of the rat ribosomal RNA gene in Chinese hamster ovary cells [59]. Although the USR activation domain of USF-2 is important for initiation in other promoter contexts [41], the USR (lacking in the dominant negative pXS1) appears to contribute predominantly to enhancer activity through E-box A, rather than the putative 'initiator' activity through motif E. However, there also appears to be co-operation or synergy between E-boxes. The binding function of USF-2 is clearly essential for initiation from the pAVP94 construct. It may be that interaction with basal transcription machinery plays an important role in activation or initiation, and that any USF dimer bound to motif E is sufficient for this to occur.

In conclusion, we have described a complex role of USF in regulating vasopressin expression in lung cancer, functioning through multiple motifs as both an enhancer (through the -147 E-box A) and perhaps as initiator (through the novel non-canonical CACGGG -42 motif E). Our data imply that bHLH factors play a major role in overcoming the threshold of repression seen in the normal state in a NSCLC model and are likely to play a part in more complex mechanisms operating in SCLC and other tissue- or stress-specific situations where vasopressin is transcribed. However, the complex interaction between bHLH factors, NRSF isoforms and other factors involved in regulating the expression of the endogenous chromatin-associated vasopressin gene remains to be fully elucidated.

This work was supported by the Association for International Cancer Research, St Andrews, Scotland (to J.M.C.), Cancer Research UK (to P.J.W.), the Medical Research Council and the Wellcome Trust (to J.P.Q.). We thank Dr M. Sawadogo (University of Texas, MD Anderson Cancer Centre, Houston, TX, U.S.A.) for the USF expression constructs.

REFERENCES

1 Ocejo-Garcia, M., Ahmed, S. I., Coulson, J. M. and Woll, P. J. (2001) Use of RT-PCR to detect co-expression of neuropeptides and their receptors in lung cancer. Lung Cancer **33**, 1–9
2 Coulson, J. M., Ocejo-Garcia, M. and Woll, P. J. (2003) The neuroendocrine phenotype of small cell lung cancer. Methods Mol. Med. **74**, 61–73
3 Ahmed, S. I., Thompson, J., Coulson, J. M. and Woll, P. J. (2000) Studies on the expression of endothelin, its receptor subtypes, and converting enzymes in lung cancer and in human bronchial epithelium. Am. J. Respir. Cell Mol. Biol. **22**, 422–431
4 Coulson, J. M., Stanley, J. and Woll, P. J. (1999) Tumour-specific arginine vasopressin promoter activation in small-cell lung cancer. Br. J. Cancer **80**, 1935–1944
5 North, W. G. (2000) Gene regulation of vasopressin and vasopressin receptors in cancer. Exp. Physiol. **85**, 27S–40S

6 Friedmann, A. S., Memoli, V. A. and North, W. G. (1993) Vasopressin and oxytocin production by non-neuroendocrine lung carcinomas – an apparent low incidence of gene-expression. Cancer Lett. **75**, 79–85

7 Burbach, J. P., Luckman, S. M., Murphy, D. and Gainer, H. (2001) Gene regulation in the magnocellular hypothalamo-neurohypophysial system. Physiol. Rev. **81**, 1197–1267

8 Coulson, J. M. (2002) Positive and negative regulators of the vasopressin gene promoter in small cell lung cancer. Progr. Brain Res. **139**, 329–343

9 Coulson, J. M., Fiskerstrand, C. E., Woll, P. J. and Quinn, J. P. (1999) Arginine vasopressin promoter regulation is mediated by a neuron-restrictive silencer element in small cell lung cancer. Cancer Res. **59**, 5123–5127

10 Coulson, J. M., Edgson, J. L., Woll, P. J. and Quinn, J. P. (2000) A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. Cancer Res. **60**, 1840–1844

11 Grace, C. O., Fink, G. and Quinn, J. P. (1999) Characterization of potential regulatory elements within the rat arginine vasopressin proximal promoter. Neuropeptides **33**, 81–90

12 Coulson, J. M., Fiskerstrand, C. E., Woll, P. J. and Quinn, J. P. (1999) E-box motifs within the human vasopressin gene promoter contribute to a major enhancer in small-cell lung cancer. Biochem. J. **344**, 961–970

13 Ball, D. W., Compton, D., Nelkin, B. D., Baylin, S. B. and Debustros, A. (1992) Human calcitonin gene-regulation by helix-loop-helix recognition sequences. Nucleic Acids Res. **20**, 117–123

14 Therrien, M. and Drouin, J. (1993) Cell-specific helix-loop-helix factor required for pituitary expression of the proopiomelanocortin gene. Mol. Cell. Biol. **13**, 2342–2353

15 Paterson, J. M., Morrison, C. F., Mendelson, S. C., McAllister, J. and Quinn, J. P. (1995) An upstream stimulatory factor (USF) binding motif is critical for rat preprotachykinin-A promoter activity in PC12 cells. Biochem. J. **310**, 401–406

16 Paterson, J. M., Morrison, C. F., Dobson, S. P., McAllister, J. and Quinn, J. P. (1995) Characterization of a functional E-box motif in the proximal rat preprotachykinin-A promoter. Neurosci. Lett. **191**, 185–188

17 van Waardenburg, R. C., Meijer, C., Pinto-Sietsma, S. J., de Vries, E. G., Timens, W. and Mulder, N. M. (1998) Effects of c-myc oncogene modulation on differentiation of human small cell lung carcinoma cell lines. Anticancer Res. **18**, 91–95

18 Chen, H., Biel, M. A., Borges, M. W., Thiagalingam, A., Nelkin, B. D., Baylin, S. B. and Ball, D. W. (1997) Tissue-specific expression of human achaete-scute homologue-1 in neuroendocrine tumors: transcriptional regulation by dual inhibitory regions. Cell Growth Differ. **8**, 677–686

19 Ito, T., Udaka, N., Ikeda, M., Yazawa, T., Kageyama, R. and Kitamura, H. (2001) Significance of proneural basic helix-loop-helix transcription factors in neuroendocrine differentiation of fetal lung epithelial cells and lung carcinoma cells. Histol. Histopathol. **16**, 335–343

20 Acampora, D., Postiglione, M. P., Avantaggiato, V., Di Bonito, M., Vaccarino, F. M., Michaud, J. and Simeone, A. (1999) Progressive impairment of developing neuroendocrine cell lineages in the hypothalamus of mice lacking the orthopedia gene. Genes Dev. **13**, 2787–2800

21 Hosoya, T., Oda, Y., Takahashi, S., Morita, M., Kawauchi, S., Ema, M., Yamamoto, M. and Fujii-Kuriyama, Y. (2001) Defective development of secretory neurones in the hypothalamus of arnt2-knockout mice. Genes Cells **6**, 361–374

22 Jin, X. W., Shearman, L. P., Weaver, D. R., Zylka, M. J., DeVries, G. J. and Reppert, S. M. (1999) A molecular mechanism regulating rhythmic output from the suprachiasmatic circadian clock. Cell **96**, 57–68

23 Sirito, M., Lin, Q., Maity, T. and Sawadogo, M. (1994) Ubiquitous expression of the 43-kda and 44-kda forms of transcription factor USF in mammalian-cells. Nucleic Acids Res. **22**, 427–433

24 Sirito, M., Lin, Q., Deng, J. M., Behringer, R. R. and Sawadogo, M. (1998) Overlapping roles and asymmetrical cross-regulation of the USF proteins in mice. Proc. Natl. Acad. Sci. U.S.A. **95**, 3758–3763

25 Sirito, M., Walker, S., Lin, Q., Kozlowski, M. T., Klein, W. H. and Sawadogo, M. (1992) Members of the USF family of helix-loop-helix proteins bind DNA as homo- as well as heterodimers. Gene Expr. **2**, 231–240

26 Ribeiro, A., Pastier, D., Kardassis, D., Chambaz, J. and Cardot, P. (1999) Cooperative binding of upstream stimulatory factor and hepatic nuclear factor 4 drives the transcription of the human apolipoprotein A-II gene. J. Biol. Chem. **274**, 1216–1225

27 Greenall, A., Willingham, N., Cheung, E., Boam, D. S. and Sharrocks, A. D. (2001) DNA binding by the ETS-domain transcription factor PEA3 is regulated by intramolecular and intermolecular protein-protein interactions. J. Biol. Chem. **276**, 16207–16215

28 Heckert, L. L. (2001) Activation of the rat follicle-stimulating hormone receptor promoter by steroidogenic factor 1 is blocked by protein kinase a and requires upstream stimulatory factor binding to a proximal E box element. Mol. Endocrinol. **15**, 704–715

29 Ismail, P. M., Lu, T. and Sawadogo, M. (1999) Loss of USF transcriptional activity in breast cancer cell lines. Oncogene **18**, 5582–5591

30 Qyang, Y. B., Luo, X., Lu, T., Ismail, P. M., Krylov, D., Vinson, C. and Sawadogo, M. (1999) Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation. Mol. Cell. Biol. **19**, 1508–1517

31 Viollet, B., Lefrancois-Martinez, A. M., Henrion, A., Kahn, A., Raymondjean, M. and Martinez, A. (1996) Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms. J. Biol. Chem. **271**, 1405–1415

32 Takahashi, K., Nishiyama, C., Okumura, K., Ra, C., Ohtake, Y. and Yokota, T. (2001) Molecular cloning of rat USF2 cDNA and characterization of splicing variants. Biosci. Biotechnol. Biochem. **65**, 56–62

33 Samoylenko, A., Roth, U., Jungermann, K. and Kietzmann, T. (2001) The upstream stimulatory factor-2a inhibits plasminogen activator inhibitor-1 gene expression by binding to a promoter element adjacent to the hypoxia-inducible factor-1 binding site. Blood **97**, 2657–2666

34 Gao, E. W., Wang, Y., Alcorn, J. L. and Mendelson, C. R. (1997) The basic helix-loop-helix-zipper transcription factor USF1 regulates expression of the surfactant protein-a gene. J. Biol. Chem. **272**, 23398–23406

35 Nicolas, G., Bennoun, M., Devaux, I., Beaumont, C., Grandchamp, B., Kahn, A. and Vaulont, S. (2001) Lack of hepcidin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (USF2) knockout mice. Proc. Natl. Acad. Sci. U.S.A. **98**, 8780–8785

36 Harris, V. K., Coticchia, C. M., List, H. J., Wellstein, A. and Riegel, A. T. (2000) Mitogen-induced expression of the fibroblast growth factor-binding protein is transcriptionally repressed through a non-canonical E-box element. J. Biol. Chem. **275**, 28539–28548

37 Virolle, T., Coraux, C., Ferrigno, O., Caillleteau, L., Ortonne, J. P., Pognonec, P. and Aberdam, D. (2002) Binding of USF to a non-canonical E-box following stress results in a cell-specific derepression of the lama3 gene. Nucleic Acids Res. **30**, 1789–1798

38 Roy, A. L., Meisterernst, M., Pognonec, P. and Roeder, R. G. (1991) Cooperative interaction of an initiator-binding transcription initiation-factor and the helix loop helix activator USF. Nature (London) **354**, 245–248

39 Terasaki, T., Matsuno, Y., Shimosato, Y., Yamaguchi, K., Ichinose, H., Nagatsu, T. and Kato, K. (1994) Establishment of a human small-cell lung-cancer cell-line producing a large amount of anti-diuretic hormone. Jpn. J. Cancer Res. **85**, 718–722

40 Dignam, J. D., Martin, P. L., Shastry, B. S. and Roeder, R. G. (1983) Eukaryotic gene transcription with purified components. Methods Enzymol. **101**, 582–598

41 Luo, X. and Sawadogo, M. (1996) Functional domains of the transcription factor USF2: atypical nuclear localization signals and context-dependent transcriptional activation domains. Mol. Cell. Biol. **16**, 1367–1375

42 Chen, Y. H., Layne, M. D., Watanabe, M., Yet, S. F. and Perrella, M. A. (2001) Upstream stimulatory factors regulate aortic preferentially expressed gene-1 expression in vascular smooth muscle cells. J. Biol. Chem. **276**, 47658–47663

43 Gobin, S. J., Montagne, L., Van Zutphen, M., Van Der Valk, P., Van Den Elsen, P. J. and De Groot, C. J. (2001) Upregulation of transcription factors controlling MHC expression in multiple sclerosis lesions. Glia **36**, 68–77

44 Groenen, P. M. A., Garcia, E., Debeer, P., Devriendt, K., Fryns, J. P. and VandeVen, W. J. M. (1996) Structure, sequence, and chromosome 19 localization of human USF2 and its rearrangement in a patient with multicystic renal dysplasia. Genomics **38**, 141–148

45 Cheung, E., Mayr, P., Coda-Zabetta, F., Woodman, P. G. and Boam, D. S. W. (1999) DNA-binding activity of the transcription factor upstream stimulatory factor 1 (USF-1) is regulated by cyclin-dependent phosphorylation. Biochem. J. **344**, 145–152

46 Luo, X. and Sawadogo, M. (1996) Antiproliferative properties of the USF family of helix-loop-helix transcription factors. Proc. Natl. Acad. Sci. U.S.A. **93**, 1308–1313

47 Sommer, A., Burkhardt, H., Keyse, S. M. and Luscher, B. (2000) Synergistic activation of the mkp-1 gene by protein kinase A signaling and USF, but not c-Myc. FEBS Lett. **474**, 146–150

48 Jaiswal, A. S. and Narayan, S. (2001) Upstream stimulating factor-1 (USF1) and USF2 bind to and activate the promoter of the adenomatous polyposis coli (APC) tumor suppressor gene. J. Cell Biochem. **81**, 262–277

49 Kaytor, E. N., Qian, J., Towle, H. C. and Olson, L. K. (2000) An indirect role for upstream stimulatory factor in glucose-mediated induction of pyruvate kinase and S14 gene expression. Mol. Cell. Biochem. **210**, 13–21

50 Smale, S. T. and Baltimore, D. (1989) The initiator as a transcription control element. Cell **57**, 103–113

51 Breen, G. A. M. and Jordan, E. M. (1999) Transcriptional activation of the F1F0 ATP synthase α-subunit initiator element by USF2 is mediated by p300. Biochim. Biophys. Acta **1428**, 169–176

52 Li, N. and Seetharam, B. (1998) A 69-base pair fragment derived from human transcobalamin II promoter is sufficient for high bidirectional activity in the absence of a TATA box and an initiator element in transfected cells – role of an E box in transcriptional activity. J. Biol. Chem. **273**, 28170–28177

53 Erickson, R. H., Lai, R. S., Lotterman, C. D. and Kim, Y. S. (2000) Identification of upstream stimulatory factor as an activator of the human dipeptidyl peptidase IV gene in Caco-2 cells. Gene **258**, 77–84

54 Sausville, E., Carney, D. and Battey, J. (1985) The human vasopressin gene is linked to the oxytocin gene and is selectively expressed in a cultured lung cancer cell line. J. Biol. Chem. **260**, 10236–10241
55 Ho, M. Y. and Murphy, D. (2002) The vasopressin gene non-canonical Hogness box: effect on protein binding and promoter function. Mol. Cell. Endocrinol. **186**, 17–25
56 Hakimi, M. A., Bochar, D. A., Chenoweth, J., Lane, W. S., Mandel, G. and Shiekhattar, R. (2002) A core-braf35 complex containing histone deacetylase mediates repression of neuronal-specific genes. Proc. Natl. Acad. Sci. U.S.A. **99**, 7420–7425
57 Ballas, N., Battaglioli, E., Atouf, F., Andres, M. E., Chenoweth, J., Anderson, M. E., Burger, C., Moniwa, M., Davie, J. R., Bowers, W. J. et al. (2001) Regulation of neuronal traits by a novel transcriptional complex. Neuron **31**, 353–365
58 Workman, J. L., Roeder, R. G. and Kingston, R. E. (1990) An upstream transcription factor, USF (MLTF), facilitates the formation of preinitiation complexes during *in vitro* chromatin assembly. EMBO J. **9**, 1299–1308
59 Ghosh, A. K., Datta, P. K. and Jacob, S. T. (1997) The dual role of helix-loop-helix-zipper protein USF in ribosomal RNA gene transcription *in vivo*. Oncogene **14**, 589–594

Received 26 July 2002/22 October 2002; accepted 28 October 2002
Published as BJ Immediate Publication 28 October 2002, DOI 10.1042/BJ20021176