MOLECULAR AND CELLULAR BIOLOGY, Jan. 1995, p. 102–111 Vol. 15, No. 1
0270-7306/95/$04.00+0
Copyright © 1995, American Society for Microbiology

Four Structurally Distinct, Non-DNA-Binding Subunits of Human Nuclear Respiratory Factor 2 Share a Conserved Transcriptional Activation Domain

SAJIV GUGNEJA, JOSEPH V. VIRBASIUS,† AND RICHARD C. SCARPULLA*

Department of Cell and Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611

Received 9 August 1994/Returned for modification 19 September 1994/Accepted 26 September 1994

Nuclear respiratory factor 2 (NRF-2) was previously purified to near homogeneity from HeLa cells on the basis of its ability to bind tandem recognition sites in the rat cytochrome oxidase subunit IV (RCO4) promoter. It consisted of five subunits, α, β_1, β_2, γ_1, and γ_2. Sequencing of tryptic peptides from α and from mixtures of the two β or two γ subunits revealed sequence identities with subunits of the mouse GA-binding protein (GABP), a ubiquitously expressed ETS domain activator composed of three subunits, α, β_1, and β_2. To understand the precise relationship between NRF-2 and GABP, cDNAs for all five NRF-2 subunits have now been cloned and their products have been overexpressed. The results establish that the two additional NRF-2 subunits are molecular variants that differ from GABP β_1 and β_2 by having a 12-amino-acid insertion containing two serine doublets. PCR and RNase protection assays show that mRNAs for these variants are expressed in the human but not the rodent cells and tissues examined. The insertion did not alter the ability of the β and γ subunits to associate with α, the DNA-binding subunit, nor did it affect the ability of NRF-2 β_1 or β_2 to direct high-affinity binding of α to tandem sites in the RCO4 promoter. In addition, the four NRF-2 β and γ subunits were equally proficient in activating transcription in transfected cells when fused to a GAL4 DNA-binding domain. The domain responsible for this transcriptional activation was localized by deletion mapping to a region of approximately 70 amino acids that is conserved in all four NRF-2 β and γ subunits. The repeated glutamine-containing hydrophobic clusters within this region bear a strong resemblance to those recently implicated in protein-protein interactions within the transcriptional apparatus.

Nuclear respiratory factors 1 and 2 (NRF-1 and NRF-2) are thought to participate in the nuclear control of mitochondrial function in mammalian cells (11, 31–34). NRF-2 was originally identified as a DNA-binding protein required for transcriptional activation of the rat cytochrome oxidase subunit IV (RCO4) gene (32). Two tandem NRF-2 recognition sites overlapping the transcription start site region behaved synergistically in maximizing promoter activity (32). The NRF-2 binding sites had the GGAA/T motif characteristic of the binding site for ETS domain transcription factors (36). This family of activators consists of both lymphoid cell-specific and ubiquitously expressed members (for reviews, see references 16, 19, and 36). Subsequently, four tandemly arranged NRF-2 recognition sites were also found to be essential for maximal expression of the mouse cytochrome oxidase subunit Vb (MCO5b) promoter (34), suggesting that NRF-2 participates in the coordinate expression of nuclearly encoded cytochrome oxidase subunits. In addition, a single NRF-2 site contributes to the transcriptional expression of mitochondrial transcription factor A (33), a nuclearly encoded factor required for mitochondrial gene transcription (12, 21, 27). These observations led to the hypothesis that NRF-2 may help coordinate the expression of respiratory chain subunits with components of the mitochondrial transcription machinery (33).

NRF-2 was purified to near homogeneity and found to have five subunits ranging in mass from 56 to 39 kDa (34). Only the 56-kDa α subunit was able to bind DNA through the GGAA core motif (34). The others, designated β_1, β_2, γ_1, and γ_2, could associate with α to form distinct DNA-protein complexes but could not themselves bind the NRF-2 recognition sites. Moreover, the β but not the γ subunits were able to confer high-affinity binding to the tandemly arranged NRF-2 sites in the RCO4 promoter (34). This behavior was reminiscent of that described for the multisubunit ETS domain transcription factor GA-binding protein (GABP) (29). GABP α is the DNA-binding subunit, while GABP β_1 and β_2 complex with α but alone do not bind DNA (17, 29). The β_1 subunit differs from β_2 in that it can homodimerize in solution and can increase the affinity of binding to closely spaced tandem sites within the herpes simplex virus ICP4 gene. This difference is associated with the presence of alternative carboxyl-terminal domains specific to each subunit. Sequencing of tryptic peptides derived from purified NRF-2 subunits revealed a high degree of sequence similarity with the GABP α and β subunits (34).

Although NRF-2 and GABP share DNA-binding properties and regions of sequence similarity, their differences in subunit composition raised the question of whether they are identical or related members of the ETS domain family. Factors that are immunologically related to GABP have been implicated in the expression of a number of genes (2, 8, 13, 20, 23). However, determination of whether these factors have common epitopes or are identical awaits their purification and structural characterization. Recently, human cDNA clones for the adenovirus early region 4 transcription factor 1 (E4TF1), a protein originally described as a multisubunit ETS domain transcriptional activator of the adenovirus E4 gene (38), have been isolated (37). The E4TF1 subunits display extensive sequence identity

* Corresponding author. Mailing address: Department of Cell and Molecular Biology, Northwestern University Medical School, 303 East Chicago Ave., Chicago, IL 60611. Phone: (312) 503-2946. Fax: (312) 503-0798.

† Present address: Program in Molecular Medicine, University of Massachusetts Medical Center, Worcester, MA 01605.

along their entire lengths with the mouse GABP α, β_1, and β_2 subunits (37). To determine the precise relationship of NRF-2 to GABP and E4TF1, we isolated cDNAs encoding NRF-2 subunits and compare the DNA-binding properties of the recombinant protein with those previously described for the native HeLa protein (34). In addition, we establish that the non-DNA-binding subunits share a transcriptional activation domain consisting of repeated hydrophobic glutamine-containing clusters spanning approximately 50 amino acid residues. Similar hydrophobic repeats have been recently associated with transcriptional activation by VP16 (5) and Sp1 (14).

MATERIALS AND METHODS

Cloning of NRF-2 cDNAs. Degenerate oligonucleotide probes for cDNA library screening were based on the sequence of peptides α94-2 and β/γ80 (34) as follows:

```
α94-2,  CAT GGG/C ATT GCC CAG CCA/T GTG/C ACG/A GCA/T GTG GCA/C CTG GC
        His  Gly    Ile  Ala  Gln  Pro    Val    Thr    Ala    Val  Ala    Leu  Ala
β/γ80,  CAT/C TGG GCA/C ACA GAG/A CAT/C AAT/C CAT/C CAG/A GAG GTG GTG GA
        His    Trp  Ala    Thr  Glu    His    Asn    His    Gln    Glu  Val  Val  Asp
```

Two HeLa cDNA libraries in λgt10 (Clontech) or λZapII (Stratagene) were screened for NRF-2 cDNAs by hybridization to oligonucleotides labeled by phosphorylation with [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs). Hybridization was carried out in 6× SSC (1× SSC is 150 mM sodium chloride plus 15 mM sodium citrate)–1× Denhardt's solution (7)–0.05% sodium pyrophosphate–100 μg of sonicated calf thymus DNA per ml at 45°C. After overnight hybridization, membranes were washed three times in 2× SSC–0.1% sodium dodecyl sulfate (SDS) for 20 min at 45°C and twice for 20 min at 55°C. Phage DNAs from positive cDNA clones detected by autoradiography were digested with *Xba*I and *Xho*I (λZapII library) or *Eco*RI (λgt10 library) and cloned into pGEM7Zf(+) (Promega) and M13mp18/mp19 for restriction analysis and DNA sequencing (25). The 5′-most *Eco*RI-*Sac*I fragment from a NRF-2β clone was used to reprobe the λgt10 library to obtain additional cDNAs. Cloned cDNAs containing intact coding regions for NRF-2 α and NRF-2 β_1 were recovered from the libraries. Coding regions for the remaining three NRF-2 subunits (β_2, γ_1, and γ_2) were reconstructed from overlapping partial cDNAs by using an internal *Msc*I site to join the appropriate 5′ and 3′ portions.

Construction of bacterial expression vectors. The NRF-2 α subunit was cloned into the bacterial expression vector pET3D (28) by converting the sequence at the natural initiator ATG codon to an *Nco*I site by PCR-directed mutagenesis using the oligonucleotide 5′-CTCCAGCCATGGCTAAAAGAG-3′ and an antisense primer complementary to sequences downstream from a natural *Eco*RI site. The resulting *Nco*I-*Eco*RI fragment was ligated with an *Eco*RI-*Bam*HI fragment produced by addition of a *Bam*HI linker at an *Ssp*I site 135 bp downstream from the NRF-2 α termination codon into the *Nco*I and *Bam*HI sites of pET3D. The NRF-2 β_1 coding region was cloned into the *Nde*I and *Bam*HI sites of expression vector pET3A (28) by ligating a synthetic *Nde*I-*Hin*dIII adaptor (5′-TATGTCCCTGGTAGATTTGGGAAAGA-3′ and 5′-AGCTTCTTTCCC AAATCTACCAGGGACA-3′) containing coding sequences of NRF-2 β_1 to the beginning of the ninth codon to a *Hin*dIII-*Bam*HI cDNA fragment produced by addition of a *Bam*HI linker at the first *Dpn*I site downstream from the termination codon. To generate bacterial expression vectors for the remaining subunits, a *Sac*I-*Bam*HI fragment (containing cDNA sequences from 91 bases downstream from the initiator ATG to the *Dpn*I site) from the β_1 expression vector described above was replaced with equivalent *Sac*I-*Bam*HI fragments from the remaining subunit clones. For β_2, the replacement fragment was in two pieces (*Sac*I-*Stu*I and *Stu*I-*Bam*HI) because of an additional *Sac*I site; for the γ clones, the downstream *Bam*HI site was generated from an *Eco*RV site in the phage polylinker.

Reverse transcriptase-mediated PCR. The following oligonucleotides were used as primers: 5′-CCTGGAGGGGTGGTGAACCT-3′ (primer I), 5′-TTCT GTCTGTAGGCCTCTGC-3′ (primer B), and 5′-GACTGCGGCAAAGCACA CCG-3′ (primer G). Primer I is in the sense orientation and begins 558 bp downstream from the ATG initiator in NRF-2 β (25 bp 5′ to the 36-bp insertion in NRF-2 β_1). NRF-2 β and γ diverge at their carboxyl-terminal regions beginning 1,036 bp from the ATG. Primers B and G are in the antisense orientation, are specific to NRF-2 β and γ, respectively, and begin 87 bp (primer B) and 17 bp (primer G) from the point of sequence divergence.

Total RNA (2 μg) or poly(A)$^+$ RNA (0.4 μg) from each source was reverse transcribed with avian myeloblastosis virus reverse transcriptase (Promega) according to the manufacturer's protocol, using primer B or primer G. One-tenth of the reverse transcriptase reaction was subsequently used for PCR. In addition to the cDNA, each PCR mixture contained 5 μl of 10× buffer (100 mM Tris-HCl [pH 8.3], 500 mM KCl, 15 mM $MgCl_2$, 0.01% [wt/vol] gelatin), 1 μl of 10 mM deoxynucleoside triphosphates, 10 pmol of primer I, 10 pmol of primer B or G, and 2.5 U of *Taq* polymerase (Perkin-Elmer Cetus) in a total volume of 50 μl. Thirty-five cycles consisting of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C were performed following a denaturation step of 3 min at 94°C in a Perkin-Elmer Cetus model 480 Thermal Cycler. Products in a 20-μl sample were then analyzed on a 4.5% acrylamide gel and visualized by ethidium bromide staining.

RNase protection analysis. The 162-bp *Apo*I-*Pst*I fragment (bases 601 to 762 from the starting ATG in NRF-2 β_1) was cloned into the *Eco*RI and *Nsi*I sites in pGEM7Zf(+). The plasmid was digested with *Pvu*II, and the antisense riboprobe was generated by using SP6 polymerase (Promega) in the presence of [α-^{32}P] CTP (Amersham) as instructed by the manufacturer. Yeast RNA was used to adjust each RNA sample to 40 μg and also used as the negative control. Positive control RNAs were transcribed in vitro from NRF-2 β_1 and β_2 cDNA templates with SP6 or T7 polymerase. Total HeLa RNA (40 μg) or human testis poly(A)$^+$ RNA (4 μg) was hybridized overnight with 2×10^5 to 5×10^5 cpm of probe at 42°C in 80% formamide–40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.4)–400 mM sodium acetate–1 mM EDTA. Following hybridization, 300 μl of digestion buffer (100 mM Tris-HCl [pH 7.5], 200 mM sodium acetate, 5 mM EDTA) containing 4 μg of RNase A per ml and 0.2 μg of RNase T_1 per ml was added, and the sample was incubated at room temperature for 1 h. The reaction was terminated by addition of 20 μl of 10% SDS and 20 μg of proteinase K, incubation at 37°C for 15 min, and extraction with 1 volume of phenol-chloroform-isoamyl alcohol (25:24:1). The RNA was precipitated with 10 μg of *Escherichia coli* tRNA as the carrier and 1 ml of ethanol. The protected fragments were resuspended in loading dye, separated on an 8% polyacrylamide–8 M urea gel, and analyzed by autoradiography.

Purification of bacterially expressed NRF-2. NRF-2 α subunit expressed following induction with isopropyl-β-D-thiogalactopyranoside (IPTG) was purified from lysates prepared and precipitated by the method of Pognonec et al. (22). The resulting ammonium sulfate precipitate was dissolved in TM buffer (25 mM Tris-HCl [pH 7.9], 6.25 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM dithiothreitol, 50 mM KCl) and loaded onto an Econo-Pac Q cartridge (Bio-Rad Laboratories) equilibrated in TM buffer. After being washed in the same buffer, proteins were eluted with a 50 to 500 mM KCl gradient in TM buffer. Fractions containing NRF-2 α were identified by Coomassie blue staining of SDS-gels and were pooled, concentrated, and adjusted to 50 mM KCl by centrifugal ultrafiltration. NRF-2 β and γ subunits induced by IPTG were purified as described for GABP (29) except that following dialysis against 25 mM Tris-HCl (pH 8.0)–75 mM NaCl–0.75 mM EDTA–10% glycerol–1 mM dithiothreitol and centrifugation, the supernatant was loaded onto an Econo-Pac Q cartridge, and the remainder of the purification carried out as with NRF-2 α.

DNA binding assays. Synthetic oligonucleotides with previously characterized NRF-2 recognition sites (32–34) used in this study were as follows:

```
RCO4 (+13 to +36),       5′-GATCCGGGACCCGCTCTTCCGGTCGCGAA-3′
                              3′-GCCCTGGGCGAGAAGGCCAGCGCTTTCGA-5′
βATPS (+582 to +605),    5′-GATCCGTGCGTTGACCTTCCGGTTGAACA-3′
                              3′-GCACGCAACTGGAAGGCCAACTTGTTCGA-5′
hCC1 (−181 to −204),     5′-GATCCTCCCCACGCTCTTCGGGTTGTCGA-3′
                              3′-GAGGGGTGCGAGAAGCCCAACAGCTTCGA-5′
hmtTFA (−34 to −13),     5′-GATCCTACCGACCGGATGTTAGCAGA-3′
                              3′-GATGGCTGGCCTACAATCGTCTTCGA-5′
MCO5b (+13 to +33),      5′-GATCCTGTTCCCGGAAGTGCATCTA-3′
                              3′-GACAAGGGCCTTCACGTAGATTCGA-5′
```

Sequences were from RCO4, ATP synthase subunit β (βATPS), human cytochrome c_1 (hCC_1), human mitochondrial transcription factor A (hmtTFA), and MCO5b genes. Mobility shift assays were performed as previously described (4, 32) with RCO4 (+13 to +36) as the probe. Reactions with purified recombinant NRF-2 contained 35 ng of α (sufficient to bind 4.2 fmol of probe) and 20 to 25 ng of β (sufficient to bind 3.4 fmol of probe when associated with α) subunits.

The region of the RCO4 gene containing NRF-2 sites (−4 to +32) was synthesized in modular fashion with the following oligonucleotides:

```
        −4                                        +32
5′-AGCTTCTTGCT CTTCCGGTGCGGGACCCGC TCTTCCGGTCA-3′
    3′-AGAACGAGAAG GCCACGCCCTGGGCGAGAA GGCCAGTGCGC-5′
```
(Bold in original: GCT CTTCCGGT and GC TCTTCCGGT in the top strand; CGAGAAG GCCA and CGAGAA GGCCA in the bottom strand.)

The NRF-2 binding sites are indicated in boldface. The +21 to +31 site was mutated by substituting AA for CC at positions +27 and +28. Oligonucleotides were ligated together into the *Mlu*I and *Hin*dIII sites of pGEM7Zf(+), and DNase I footprinting was performed as previously described (11). The amount of recombinant NRF-2α that could bind 1.3 fmol of RCO4 (+13 to +36) probe in a mobility shift assay was arbitrarily designated 1× (~11 ng). The amount of β subunit added was equal to the amount required to convert all of the α subunit in the reaction to an αβ complex under the standard mobility shift assay conditions. Cellular NRF-2 was purified from HeLa nuclear extract as previously described (34).

GAL4 fusion proteins. Full-length β subunits as well as C-terminal deletions of the β_1 subunit were fused to the GAL4 DNA-binding domain [Gal 4(1–147)] by conversion of a *Hin*dIII site at codon 9 of the NRF-2 β sequence to a *Bam*HI site by linker addition. This *Bam*HI site along with an *Xba*I site derived from the polylinker in NRF-2 subunit subclones was used for cloning into the same site in the vector pSG424 (24) which contained GAL4(1–147) expressed from the simian virus 40 early promoter. In these full-length fusion proteins, the natural NRF-2 subunit termination codon was used. Most of the carboxyl-terminal deletions of the GAL4-β_1 fusion protein were generated by exonuclease III deletion from the *Xba*I site followed by *Xba*I linker addition and religation to the *Xba*I site of the vector, utilizing the universal terminator sequence downstream of the multiple cloning site in pSG424. Carboxyl-terminal deletions to amino acid 345 were produced by insertion of an *Xba*I site at that position by PCR-mediated mutagenesis, using the primer 5′-CGCTCTAGACTTCTATTTCTGC-3′. Constructions with N-terminal endpoints at amino acids, 227, 258, and 308 were generated by insertion of a *Bam*HI site at those positions, using the oligonucleotides 5′-CGGGATCCTAGCTGCCTTAG-3′, 5′-CGGGATCCTGGATGGT CCCA-3′, and 5′-CGGGATCCTAACAGTACCAG-3′, respectively. For the GAL4(1–147) NRF-2 α fusion, a *Kpn*I site was inserted at the second codon of the NRF-2 α coding region by PCR using the oligonucleotide 5′-CCGGTAC CTGACTAAAAGAGAAGC-3′. An *Xba*I site was added by linker addition to an *Msc*I site at codon 318 (immediately upstream of the ETS DNA-binding domain). The resulting *Kpn*I-*Xba*I fragment was ligated to the same sites of pSG424.

Cell culture. Growth and transfection of COS cells by the calcium phosphate method were as previously described (10). For each transfection, three 100-mm-diameter plates were each transfected with 2 μg of the reporter, G_5BCAT (18), 1 μg of the GAL4 fusion expression plasmid, and 27 μg of carrier plasmid DNA. After 48 h, cells from the triplicate plates were harvested, pooled, and divided into three equal portions. One third was used for the preparation of extracts for chloramphenicol acetyltransferase (CAT) assay as previously described (10). Another third was used for the extraction of low-molecular-weight DNA by the method of Hirt (15), which was assayed for the amount of CAT-coding plasmid DNA as previously described (9) to normalize transfection efficiency. Nuclear protein extracts were prepared from the remaining third as described previously (1) to assay the expression of the GAL4 fusion proteins by mobility shift assays using a GAL4-binding-site oligonucleotide.

Nucleotide sequence accession numbers. cDNA sequences have been submitted to GenBank under the following accession numbers: α, U13044; β_1, U13045; β_2, U13046; γ_1, U13047; and γ_2, U13048.

RESULTS

Cloning of NRF-2 subunits. The limited tryptic peptide sequencing performed previously did not rule out the possibility that NRF-2 and GABP represent different members of the ETS domain family that have regions of sequence similarity. In addition, it was unclear whether the differences in the number of subunits (five for NRF-2 versus three for GABP) resulted from posttranslational modifications or whether the additional non-DNA-binding subunits of NRF-2 represent true genetic variants encoded in the DNA. To resolve these questions, tryptic peptide sequences from the α subunit of NRF-2 and from a peptide common to β and γ subunits were used to design degenerate oligonucleotide probes for screening HeLa cDNA libraries. Overlapping NRF-2 α clones isolated with the α-peptide probe contained an open reading frame of 1,362 bp, included the ETS domain, and had 76% nucleotide sequence identity with mouse GABP α (not shown). The predicted mass of 51.4 kDa for this subunit was somewhat less than the 56 kDa estimated for the HeLa α subunit on denaturing gels (34).

A cDNA clone isolated with the NRF-2 β/γ probe contained an open reading frame of 1,185 bp with a predicted mass of 42.5 kDa (Fig. 1A, β/γ-1). This clone had 86% nucleotide sequence identity with GABP β_1 and included the β_1-specific carboxyl-terminal domain. However, it differed by having a 36-bp insertion starting 583 nucleotides from the initiator ATG. This insertion encodes 12 amino acids including two serine doublets and is not found in any of the known subunits of GABP. A second cDNA, β/γ-2, began 155 bp into the coding region of β/γ-1 and contained the same insertion but unlike β/γ-1 had a carboxyl-terminal domain homologous to GABP β_2. A third cDNA clone, β/γ-3, had an incomplete 3′ end and was clearly missing the insertion but was otherwise identical to the other two except for a synonymous A-to-G transition at Lys codon 69. These results raised the possibility that the human counterparts to GABP β_1 and β_2 each exist in two versions that differ only by the presence of the insertion thus accounting for the two additional NRF-2 subunits (Fig. 1B).

Because only two of the cDNA clones contained the insertion, both from the same library, it was important to establish its presence in both β- and γ-encoding cellular RNAs. To this end, total RNA samples from four human and two rat sources were reverse transcribed by using the NRF-2 β-specific (Fig. 1A, primer B) or γ-specific (primer G) primer. Subsequent amplification of each cDNA sample was performed with the same sense primer complementary to sequences just upstream from the 36-bp insertion (Fig. 1A, primer I). Two amplified bands of 584 and 548 bp corresponding to NRF-2 β_1 and β_2 should result from the primer I-B combination, whereas amplified bands of 514 and 478 bp corresponding to γ_1 and γ_2 should result from the primer I-G combination. Amplification of products by using the same primers on cloned cDNAs for each subunit served as controls.

In each reaction in which human RNA was used, two fragments of the expected sizes were observed with each primer combination (Fig. 2). In addition, a larger fragment was also produced (primers I and B, lanes 5 to 8; primers I and G, lanes 14 to 17). This additional fragment was determined to be a heteroduplex consisting of one strand from the β_1 (or γ_1) product and the other from the β_2 (or γ_2) product by the following criteria: (i) upon isolation and reamplification of the large fragment, all three fragments were produced (data not shown); (ii) cloning of the third fragment followed by restriction digestion with flanking restriction sites always yielded the two smaller fragments expected but not the third; and (iii) while cloned cDNAs of β_1 and β_2 would yield a single fragment when amplified individually with primers I and B (lanes 2 and 3), three fragments were produced when the cDNAs were mixed together prior to amplification (lane 4). This third fragment was the same size as the additional fragment generated via amplification from human RNA. These results were also true for γ_1 and γ_2 amplification with primers I and G (lanes 11 to 13). In contrast, the single dominant PCR product from the rat tissue RNAs (lanes 9, 10, 18, and 19) suggests that the rat has only one version of each subunit.

To confirm these results, RNase protection analysis was performed. Taking advantage of an *Apo*I restriction site within the insertion, a riboprobe which partially spanned the insertion (Fig. 1A) was hybridized to RNA from HeLa cells and human testis. In both cases, two bands were protected from RNase digestion (Fig. 3, lanes 5 and 6). These were identical to the protected bands produced when in vitro-transcribed RNA from NRF-2 β_1 or β_2 was hybridized to the probe (lanes 2 and 3). Yeast RNA, used as the negative control, yielded no protected bands (lane 4). These results provide independent confirmation that the serine-rich insertion is encoded in expressed NRF-2 β and γ transcripts. This insertion, in combination with the GABP β_1 and β_2-specific carboxyl-terminal domains, can account for the existence of the four non-DNA-binding subunits in human NRF-2 (Fig. 1B).

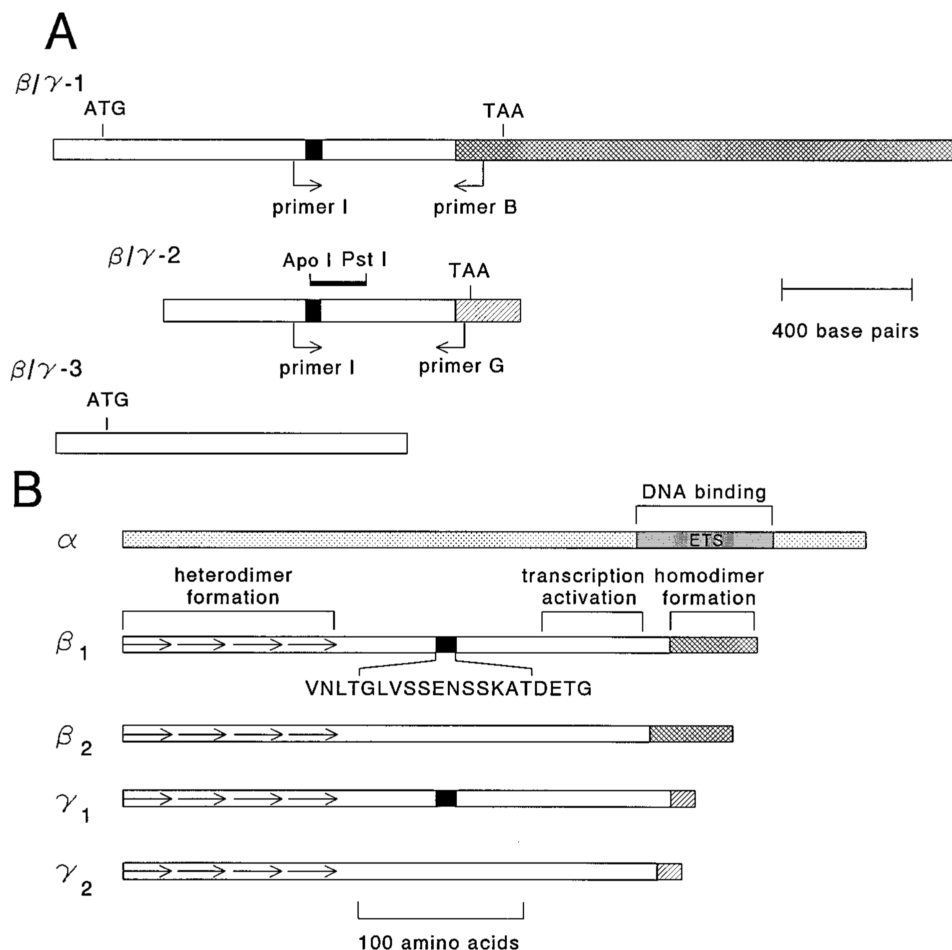

FIG. 1. Cloning and characterization of NRF-2 subunits. (A) Overlapping cDNAs encoding NRF-2 β and γ subunits were isolated from HeLa cDNA libraries by hybridization and designated β/γ-1, β/γ-2, and β/γ-3. Locations of the initiator methionine codon (ATG) and stop codon (TAA) are indicated. Arrows indicate the locations and orientations of primers used for reverse transcriptase-mediated PCR of NRF-2 β and γ mRNAs (Fig. 2). The *Apo*I-*Pst*I fragment from NRF-2 β used as the probe for the RNase protection assay (Fig. 3) is shown as a horizontal bar. (B) Summary of the structural and functional features of NRF-2 subunits from information contained in this report and from the analysis of GABP (17, 29) and E4TF1 (37). Arrows indicate the ankyrin repeats in NRF-2 β and γ subunits as originally observed in GABP β_1 and β_2 and implicated in heterodimerization with the DNA-binding α subunit. The α subunit is stippled, with the ETS domain shaded. In both panels, the NRF-2 β-specific and NRF-2 γ-specific carboxyl-terminal domains are shown by cross-hatched and hatched boxes, respectively; the serine-rich insertion in β_1 and γ_1 is indicated by a filled box, with the sequence given below. Portions of the proteins implicated in DNA binding, homodimer formation, and transactivation are also indicated.

Expression of NRF-2 subunits in *E. coli* and characterization of recombinant NRF-2. To determine whether the cDNA-encoded subunits are functionally equivalent to those purified from HeLa cells, each cDNA was expressed in *E. coli* by using the pET expression system (28). The IPTG-induced recombinant proteins were purified to near homogeneity (Fig. 4). Their molecular masses on denaturing SDS-gels (α, 58 kDa; β_1, 48 kDa; β_2, 46 kDa; γ_1, 38 kDa; γ_2, 37 kDa) were indistinguishable from those previously determined for the purified HeLa subunits (34) (Table 1). Binding of purified HeLa protein or a mixture of recombinant NRF-2 subunits to a single NRF-2 site from the RCO4 gene (RCO4 [+13 to +36]) yielded the same three DNA-protein complexes in mobility shift gel assays that are characteristic of NRF-2 binding to its recognition site (Fig. 5, lanes 1 and 2). These consist of a slower-migrating doublet comprised of αβ (upper band) and αγ (lower band) heteromers as well as a faster-migrating complex containing α alone (34). All three complexes were competed for by an excess of unlabeled oligomers of known NRF-2 recognition sites (lanes 5 to 7) but not by nonspecific oligomers containing a single G-to-C transversion in the GGAA core motif (lanes 3 and 4). As observed previously with the purified HeLa NRF-2 subunits (34), only α was able to bind DNA (lane 8); the β and γ subunits were not (lane 9). When mixed with α, the β (lanes 10 and 11) and γ (lanes 12 and 13) subunits formed heteromeric complexes corresponding to those observed with native HeLa NRF-2 (lane 1) or the mixture of five recombinant subunits (lane 2). Therefore, the recombinant and HeLa subunits are indistinguishable in molecular mass and in the ability to participate in the formation of distinct DNA-protein complexes.

We previously found that the β but not the γ subunit of NRF-2 conferred high-affinity binding to tandem recognition sites in the RCO4 promoter region (34). If the recombinant proteins described here encode these subunits, they should exhibit the same properties. Recombinant NRF-2 α and all four possible αβ and αγ complexes were used for DNase I footprinting of the two tandem NRF-2 sites (sites A and B) in the RCO4 promoter (Fig. 6A). NRF-2 α alone (lanes 3 to 6) and all αβ (lanes 7 to 14) and αγ (lanes 15 to 22) complexes protected probe DNA from cleavage by DNase I and gener-

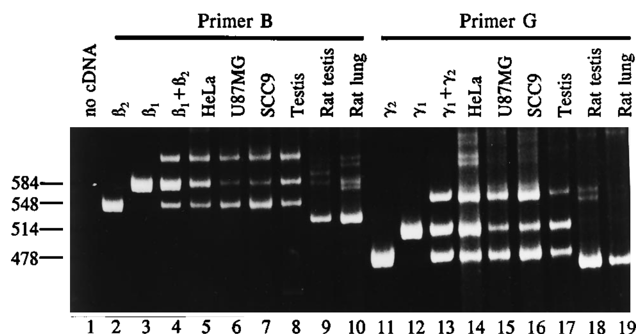

FIG. 2. Reverse transcriptase-mediated PCR amplification of NRF-2 β and γ mRNAs from human and rat cells. Total RNA (2 μg) or poly(A)$^+$ RNA (0.4 μg) was primed for reverse transcription with 10 pmol of NRF-2 β- or γ-specific primer (primer B; lanes 2 to 10, primer G; lanes 11 to 19), respectively (see Fig. 1A for primer locations). The same primers were then used in conjunction with primer I for PCR as described in Materials and Methods. For positive controls, the cDNA of each subunit cloned into pGEM7Z was amplified individually (lanes 2, 3, 11, and 12) or in pairs (lanes 4 and 13). Products were separated on a 4.5% polyacrylamide gel and visualized by ethidium bromide staining. RNA was from human tissues or cell lines except for rat testis (lanes 9 and 18) and rat lung (lanes 10 and 19). U87MG and SCC9 are human glioblastoma and squamous cell carcinoma lines, respectively. The sizes of the expected bands are indicated at the left in base pairs.

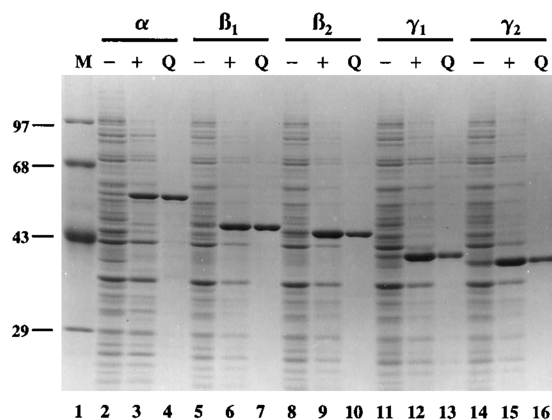

FIG. 4. Expression and purification of recombinant NRF-2 subunits from *E. coli*. cDNAs for each NRF-2 subunit were cloned into pET expression vectors and transformed into *E. coli*. Bacterial cultures were uninduced (−) or induced (+) with 0.4 mM IPTG to express the NRF-2 subunit indicated at the top. A small aliquot of bacteria from each culture was denatured by boiling in SDS buffer and loaded on an SDS–10% polyacrylamide gel. Purification of recombinant NRF-2 subunits from bacterial lysates was performed with an Econo-Pac Q cartridge (Q) as described in Materials and Methods. Gel lanes contained ~1 μg of purified protein. Proteins were visualized by FastStain (Zoion Research). Molecular mass standards in kilodaltons are indicated (lane M).

ated footprints similar to those previously observed with the native HeLa protein (32, 34). The heteromeric complexes produced more extended footprints (lanes 10, 14, 18, and 22) compared with that formed with α alone (lane 6). Binding of α to tandem sites rendered the nucleotides between the RCO4 A and B sites hypersensitive to DNase I cleavage (lanes 3 to 7). The presence of either β or γ extended the footprint, reducing the number of hypersensitive cleavages in this region (compare lane 6 with lanes 10, 14, 18, and 22). Those remaining were more enhanced with β (lanes 10 and 14) than with γ (lanes 18 and 22), possibly reflecting more distortion of the intervening DNA as a result of β-subunit interactions.

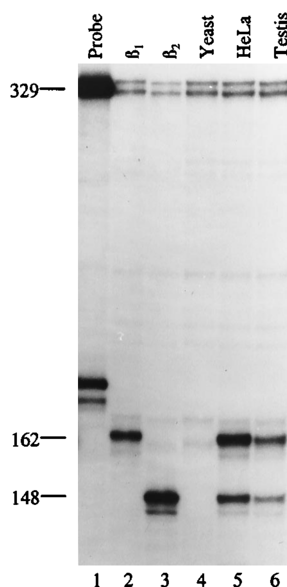

FIG. 3. Detection of the β_1- and γ_1-specific 36-nucleotide insertion in human RNA. The 162-bp *Apo*I-*Pst*I fragment from NRF-2 β_1 (indicated in Fig. 1) was used as the template to generate the antisense riboprobe as described in Materials and Methods. HeLa total RNA (40 μg; lane 5) or human testis poly(A)$^+$ RNA (4 μg; lane 6) was hybridized overnight with the radiolabeled probe. RNA transcribed in vitro from cDNA clones containing the full coding region of NRF-2 β_1 (lane 2) or β_2 (lane 3) was used as the positive control, and yeast RNA was used as the negative control (lane 4). Following hybridization, samples were digested with RNases A and T_1 for 1 h at room temperature. After inactivation of RNases, protected fragments were recovered by ethanol precipitation, analyzed on a denaturing 8% polyacrylamide–urea gel, and visualized by autoradiography. The sizes of the probe and the protected bands are indicated at left in base pairs.

Complete protection from DNase I cleavage was achieved with approximately 10-fold-less binding activity (8× versus 80×), using the $\alpha\beta_1$ (lanes 7 to 10) and $\alpha\beta_2$ (lanes 11 to 14) complexes compared with α alone (lanes 3 to 6), $\alpha\gamma_1$ (lanes 15 to 18), or $\alpha\gamma_2$ (lanes 19 to 22). This result is identical to that obtained with the purified HeLa subunits (34) and is consistent with the observed behavior of GABP β_1 and β_2 (29). By contrast, with a promoter fragment containing a mutated A site, the difference in the amount of binding activity required for full protection between the β (Fig. 6B, lanes 7 to 14)- and γ (lanes 15 to 22)-containing complexes was much less (28× versus 80×) than with intact tandem sites. Thus, as observed with native NRF-2, the enhanced binding conferred by the β subunits is lost when one site is mutated (34). The serine-rich insertion had no effect on binding when present in either β or γ, suggesting that its presence alone does not regulate the binding activity of NRF-2. Thus, the DNA-binding properties

TABLE 1. Comparison of NRF-2, GABP, and E4TF1 subunits

NRF-2		GABP		E4TF1	
Subunit	Molecular mass (kDa)[a]	Subunit	Molecular mass (kDa)	Subunit	Molecular mass (kDa)
α	56 (51.4)	α	(51.3)	E4TF1-60	(51.4)
β_1	48 (42.5)				
β_2	46 (41.3)	β_1	(41.3)	E4TF1-53	(41.3)
γ_1	41 (38.1)				
γ_2	39 (36.9)	β_2	(37.0)	E4TF1-47	(36.9)

[a] Experimentally determined molecular mass, with the value predicted from each subunit cDNA coding region in parentheses.

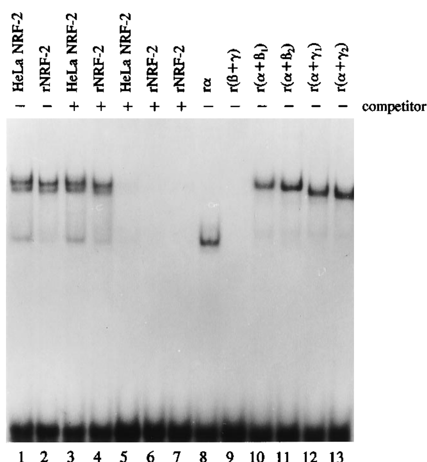

FIG. 5. Comparison of DNA-protein complexes formed with recombinant NRF-2 subunits with those formed with the purified HeLa protein. The end-labeled RCO4 (+13 to +36) oligonucleotide containing a single NRF-2 binding site was used as the probe in mobility shift assays. DNA-protein complexes formed by using affinity-purified HeLa NRF-2 are shown in lanes 1, 3, and 5. Recombinant subunits (designated r) were purified with an Econo-Pac Q cartridge as described in Materials and Methods. Recombinant NRF-2 (lanes 2, 4, 6, and 7) consisted of a mixture of ~35 ng of NRF-2 α and 5 to 7 ng each of NRF-2 β_1, β_2, γ_1, and γ_2. The same amounts of these subunits were used in reactions containing α alone (lane 8) or a mixture of β and γ subunits without α (lane 9). Amounts of β or γ subunits were increased to 20 to 25 ng when mixed individually with α (lanes 10 to 13). Unlabeled competitor oligonucleotides were hCC_1 (−181 to −204) (lanes 3 and 4), MCO5b (+13 to +33) (lane 5), hmtTFA (−34 to −13) (lane 6), and βATPS (+582 to +605) (lane 7). The hCC_1 (−181 to −204) oligomer differs from RCO4 (+13 to +36) by a single G→C transversion that eliminates NRF-2 binding and was used as a nonspecific competitor.

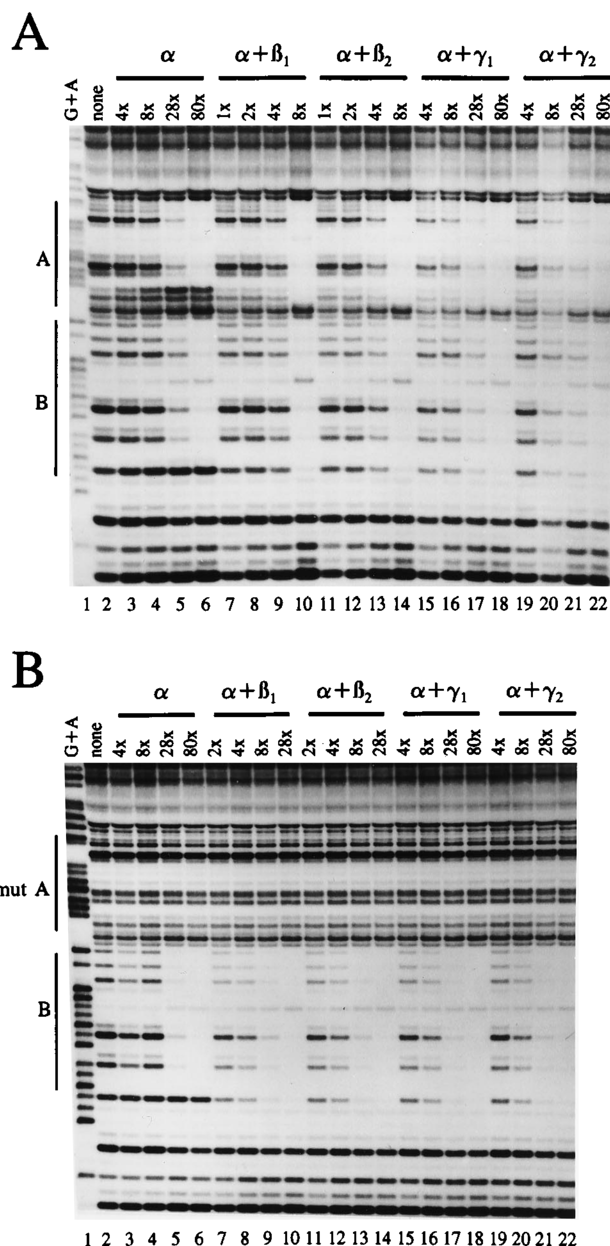

FIG. 6. DNase I footprinting of tandem binding sites by recombinant NRF-2 subunits. (A) A RCO4 promoter fragment labeled on the noncoding strand was incubated with either no protein (lane 2), different amounts of recombinant NRF-2α (lanes 3 to 6), or mixtures of α with β or γ subunits (lanes 7 to 22) and subjected to DNase I cleavage. 1× was defined as the amount of NRF-2 α required to bind ~1.3 fmol of probe containing a single NRF-2 binding site in a mobility shift assay (~11 ng of NRF-2 α). The amount of β or γ added was sufficient to displace all of the α subunit into a heteromeric complex as measured by mobility shift assay. (B) Same as panel A except that NRF-2 binding site A was inactivated (mut A) by replacing the GGAA with TTAA in the labeled RCO4 promoter fragment.

of the recombinant subunits on tandem sites are indistinguishable from those of the HeLa protein.

Transcriptional activation by NRF-2 subunits and localization of the activation domain. It was of interest to determine which of the NRF-2 subunits are involved in transcriptional activation. Since the β and γ subunits must associate with α to bind DNA, it was important to distinguish the structural features required for heteromer formation from those required for transcriptional activation. Thus, the subunits were fused to a GAL4 DNA-binding domain to eliminate the need for binding to α and allowing an independent analysis of the activation function. Because both heteromerization and DNA-binding domains are localized to the carboxyl-terminal region of α (29), the region downstream from amino acid residue 318 was not included in the GAL4 fusion construction. This fusion did contain NRF-2 α residues 2 to 318, including a region rich in acidic residues at the extreme amino terminus. As shown in Fig. 7A, the GAL4-α fusion protein containing α residues 2 to 318 failed to stimulate activity of the reporter plasmid. Additional fusions including smaller amino-terminal portions but still including the acidic region near the amino terminus also failed to stimulate transcription (results not shown), making it unlikely that activity is masked by downstream sequences.

The β or γ subunits were fused to GAL4 at amino acid 10, leaving the remainder of each intact. In contrast to α, each of the GAL4-β and -γ fusions stimulated transcription significantly (at least 100-fold compared with the activity obtained in cotransfection with the GAL4 DNA-binding domain alone), and all four demonstrated equal potency. All five of the fusion proteins were similarly expressed in cotransfected cells, as demonstrated by mobility shift assays using a labeled GAL4 binding site (Fig. 7B). Cells transfected with each of the GAL4 fusion proteins (lanes 3 to 7) contained a binding activity not present in cells transfected with the reporter plasmid alone (lane 1). The complexes produced from the fusion proteins had a slower mobility than the complex formed with extracts from cells transfected with the GAL4 DNA-binding domain alone (lane 2), indicating full-length translation of the fusion. Although it is formally possible that an activation region exists within the ETS or carboxyl-terminal region of NRF-2 α, the

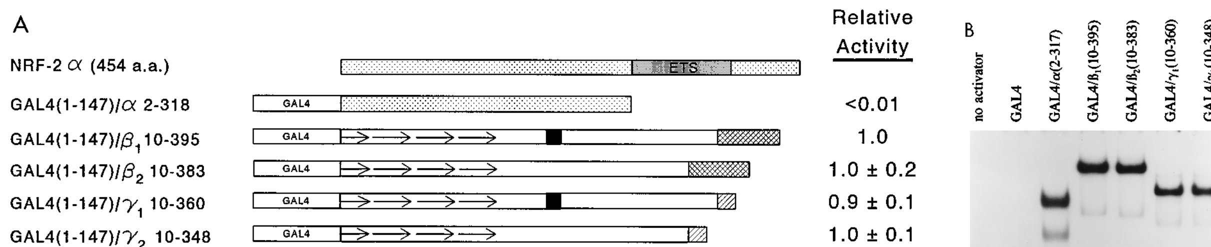

FIG. 7. Transcription activation by NRF-2 subunits. (A) NRF-2 subunits were fused to GAL4(1–147) and tested for transcriptional activation of a cotransfected reporter CAT plasmid containing five GAL4 binding sites. A diagrammatic representation of the NRF-2 α subunit (stippled) is at the top; the ETS DNA-binding domain is shaded. The ankyrin/notch repeats (arrows), the serine-rich insertion (filled box), and the carboxyl-terminal regions specific to the β (cross hatched) and γ (hatched) subunits are depicted in maps of the β- and γ-subunit fusions. The activity of the β_1 fusion [GAL4(1–147)/β_1 10–395] is defined as 1.0 and is more than 100-fold greater than the activity of the GAL4 DNA-binding domain alone. Results are given as the average ± standard deviation of at least three determinations. a.a., amino acids. (B) Detection of GAL4 fusion proteins in transfected cell extracts. Nuclear protein extracts were prepared from cells transfected with the reporter plasmid alone (lane 1), reporter plus GAL4(1–147) expression vector (lane 2), or reporter plus GAL4 fusion expression plasmids (lanes 3 to 7). Complexes formed upon incubation of these extracts with a labeled GAL4 binding site oligonucleotide were resolved on native gels and detected by autoradiography.

simplest interpretation of the results is that the β and γ subunits of NRF-2 are responsible for transcriptional activation.

Given that the four β and γ subunits were equal in the ability to activate transcription when fused to a heterologous DNA-binding domain, it was anticipated that the location of the activation domain would be within a region common to all four subunits. Fusion proteins with various deletions of β_1 linked to the GAL4 DNA-binding domain were therefore tested by cotransfection with the GAL4 reporter (Fig. 8A). Deletion from the carboxyl terminus to amino acid 345 (which precisely removes the β_1-specific region), 327, or 311 increased the transcriptional activity relative to the full-length β_1 fusion (whose activity is defined as 1.0). Further deletion to 268 markedly reduced activity, and deletion to 244 eliminated it, placing the carboxyl-terminal boundary of the activation domain in the vicinity of residue 311. Stable expression of each of the fusion proteins was verified by mobility shift assays using a GAL4 binding site (Fig. 8B). As noted above, no GAL4 binding activity was recovered from cells transfected with only the reporter plasmid (lane 1). A complex was detected in cells transfected with the GAL4 DNA-binding domain alone (lane 2). Complexes of mobility slower than GAL4(1–147) and reflective of the expected size of the construct were recovered from cells transfected with each fusion protein expression vector (lanes 2 to 13). The observed heterogeneity in the banding pattern with certain fusion proteins appears to result from their relative instability.

To further define the boundaries of the activation domain, amino-terminal truncations of β_1 were fused to GAL4 and tested in cotransfection (Fig. 8A). Deletion to residue 227 resulted in activity identical to that of the most active carboxyl-terminal deletions, indicating that neither the ankyrin/notch repeats nor the serine-rich insertion is required for transcription activation. Deletion of β_1 sequences to amino acid 258 increased activity, but further deletion to 308 abolished it completely placing the amino-terminal boundary of the activation domain between amino acids 258 and 308.

Several smaller fusions encompassing the active region defined by carboxyl-terminal deletions (β_1 227–345, β_1 258–345, and β_1 258–311) were also tested, and each was active in stimulating transcription. The smallest functionally independent region, β_1 258–311, retained about one-half of the activity of the full-length β_1 fusion, most likely because it was more unstable than the others (Fig. 8B, lane 13). A specific contribution by residues downstream from 327 is unlikely because carboxyl-terminal deletions in this region (β_1 10–345 and β_1 10–327) increased activity and β_1 308–395 was completely inactive. These results are consistent with the boundaries of an activation domain lying between residues 258 and 327, a sequence identical in the four β and γ subunits lying between the serine-rich insertion and the carboxyl-terminal domain that distinguishes β from γ. Interestingly, this domain has a number of evenly spaced clusters of hydrophobic residues containing glutamine. A similar motif is present in several activation domains of the glutamine-rich class and has recently been proposed to form an interface for interaction with the general transcription machinery (30). Therefore, each β or γ subunit of NRF-2 likely utilizes the same structural domain to activate transcription.

DISCUSSION

NRF-2 subunit composition. Both RCO4 and MCO5b promoters are unusual in that their activities are almost completely dependent on tandemly arranged repeats of the GGAA/T motif recognized by NRF-2 (32, 34). Here, we have cloned five cDNAs by using degenerate oligomers derived from NRF-2 tryptic peptides. The recombinant proteins encoded by these cDNAs display the same DNA-binding properties and structural features previously observed for the five NRF-2 subunits purified from HeLa cells. They have the same molecular masses, are expressed in human cells and tissues, and form distinct heteromeric complexes that bind DNA. These results provide compelling evidence that the cDNAs described here account for the NRF-2 subunits expressed in human cells and tissues.

The cloning of the human NRF-2 subunits has resolved the previously observed discrepancies in subunit composition between human NRF-2 and mouse GABP (17, 29, 34). The two additional NRF-2 subunits are variants of GABP β_1 and β_2 that are larger because of a 12-amino-acid insertion within their coding regions (Fig. 1B; Table 1). This conclusion is supported by the detection of four expressed RNA transcripts that encode the two possible carboxyl-terminal domains in

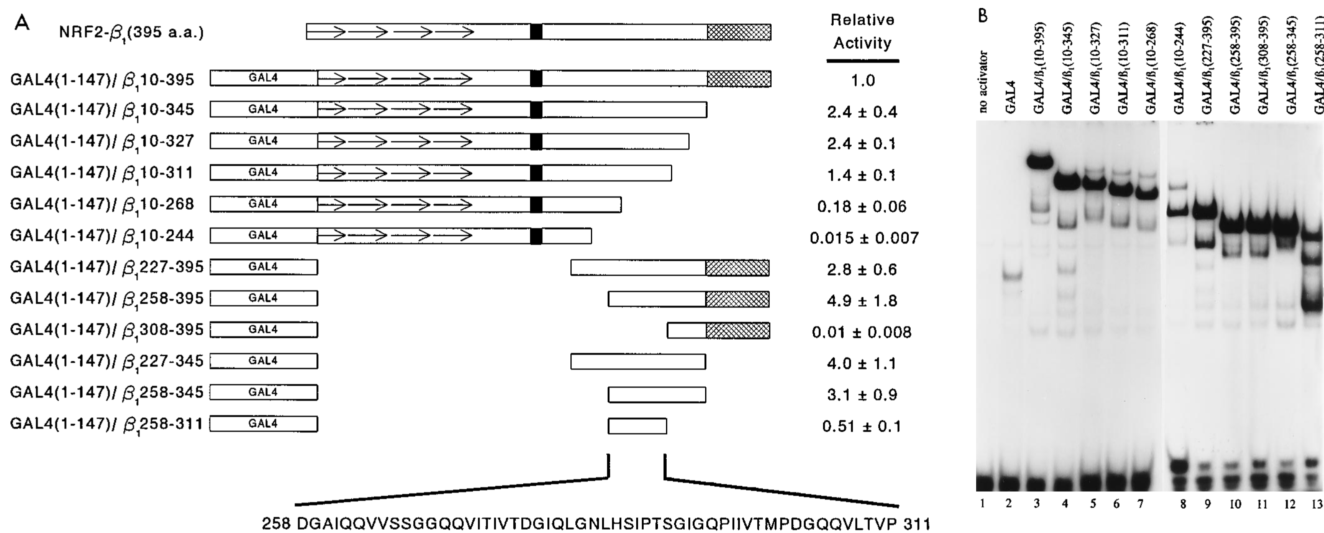

FIG. 8. Localization of the NRF-2 transcription activation domain. (A) Portions of the NRF-2 β_1 subunit were fused to GAL4(1–147) and tested for transcriptional activation of a cotransfected CAT reporter plasmid containing five GAL4 binding sites. The activity of the nearly full-length β_1 fusion [GAL4(1–147)/β_1 10–395] is defined as 1.0 and more than 100-fold greater than the activity of the GAL4 DNA-binding domain alone. The ankyrin/notch repeats (arrows), the serine-rich insertion (filled box), and the carboxyl-terminal regions specific to the β subunits (cross-hatched) are diagrammatically represented at the top. The complete sequence of the NRF-2 β_1 subunit between amino acids (a.a.) 258 and 311 (which is identical in all β and γ subunits) is given below, with repeated glutamine-containing hydrophobic clusters underlined. CAT activities relative to the nearly full-length GAL4(1–147)/β_1 10–395 are expressed as the average ± standard deviation of at least three determinations. (B) Detection of GAL4 fusion proteins in transfected cell extracts. Nuclear protein extracts were prepared from cells transfected with the reporter plasmid alone (lane 1), reporter plus GAL4(1–147) expression vector (lane 2), or reporter plus GAL4 fusion expression plasmids (lanes 3 to 13). Complexes formed upon incubation of these extracts with a labeled GAL4 binding site oligonucleotide were resolved on native gels and detected by autoradiography.

combination with or without the insertion. Transcripts containing this serine-rich insertion were observed in human but not rat cells and tissues and may be species specific. The insertion in NRF-2 β_1 or γ_1 had no discernable effect on DNA binding or on transcriptional activation. However, it may serve as the site of a posttranslational modification that may regulate NRF-2 activity during cell growth or differentiation. It is of interest in this context that at least four developmentally regulated proteins reactive to anti-GABP β antiserum have been detected in *Xenopus laevis* (20), although it remains to be determined whether any contain the insertion described here. The nucleotide sequence at the ends of the region encoding the insertion is consistent with the 5′ end being a splice acceptor and the 3′ end being a splice donor, suggesting that these variants arise via alternate splicing. This view is supported by the fact that sequences common to β/γ-3 and β/γ-1 cDNAs (Fig. 1A) are identical except for the insertion and a single transition. Moreover, recent findings indicate that the GABP β_1 and β_2 subunits are encoded by a single gene (6).

Three of the five NRF-2 subunits, NRF-2 α, β_2, and γ_2 are identical to those recently reported for human E4TF1 (Table 1) (37). This factor was originally described as a two-subunit member of the ETS family that stimulated the adenovirus early region 4 promoter and had separate subunits for DNA binding and transcription activation (38). Subsequently, three subunits, E4TF1-60, -53, and -47, which correspond to GABP α, β_1, and β_2, respectively, were cloned (Table 1) (37). Doublets of E4TF1-53 and -47 from HeLa cells were detected by Western blotting (immunoblotting) using anti-GABP β antiserum but were ascribed to posttranslational modifications (37). Our results suggest that at least part of the difference in mobility results from the presence of the 12-amino-acid insertion described here.

DNA binding. When present in heteromeric complexes with α, the β subunits of HeLa NRF-2 were both able to increase the affinity of binding to tandem recognition sites without altering binding specificity (34). The αγ complexes had a much reduced affinity for tandem sites, and both αβ and αγ displayed the same low-affinity DNA binding when one of the two tandemly arranged sites was mutated. As shown here, the recombinant subunits display the same DNA-binding properties and yield DNA-protein complexes identical in mobility to those formed with the purified HeLa subunits. The cooperativity between tandem sites mediated by the NRF-2 $\alpha\beta_1$ or $\alpha\beta_2$ complexes occurred whether the sites were 20 bp apart (center to center) as they are in the RCO4 promoter (32) or 30 bp apart as in MCO5b (sites C and D [34]). We also find there is no difference in the amount of NRF-2αβ complex required for complete DNase I protection when the RCO4 sites are 15 bp apart, indicating that rotation of the binding sites onto different faces of the DNA helix is not a critical factor (not shown). High-affinity binding of GABP $\alpha\beta_1$ complexes to the herpesvirus ICP4 promoter occurs when the sites are 6 bp apart (29). As noted previously (29), the results suggest a surprising degree of structural flexibility, leading to the subunit interactions required for high-affinity binding to tandem sites in promoters.

These observations may have important implications for the genes controlled by these factors. The level of promoter activation by GABP or NRF-2 depends upon (i) whether a GGAA/T motif is present, (ii) the number of such sites present (monomer versus tandemly repeated) and (iii) the relative levels of the various subunits that are expressed in cells. We have found that the tandem arrangement of NRF-2 sites in the RCO4 and MCO5b promoters renders their activities almost completely dependent on these sites in transfected cells (32, 34). Such a strong requirement for multiple ETS domain-binding sites was also recently observed for the β2 integrin promoter (2). Two or more such binding sites have been observed in the ribosomal protein L30 (13), folate-binding protein (23), and aldose reductase (35) genes. By contrast, we find

the mitochondrial transcription factor A promoter to be much less dependent on its single NRF-2 site (33). Single sites are also present in several other genes, including the ATP synthase β subunit (32), ribosomal protein L32 (13), and 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (8) genes.

Transcriptional activation. The ability of the NRF-2 subunits to activate transcription was assayed by fusing them to a yeast GAL4 DNA-binding domain. This allowed an analysis of transcription activation that was independent of a requirement for heteromerization with the DNA-binding α subunit. Fusion proteins with each of the four NRF-2 β and γ subunits were identical in the ability to activate the transcription of a reporter gene in transfected cells. The level of stimulation was 100-fold greater than with the GAL4 DNA-binding domain alone or a GAL4 fusion containing the α subunit that had been deleted of carboxyl-terminal sequences downstream of and including the DNA-binding ETS domain. These results are consistent with the in vitro results obtained with E4TF1 showing that E4TF1-53 (NRF-2 β_2 [Table 1]) is required for transcription (26, 38). The mapping described here localizes the activation domain to residues 258 to 327, which lie between the ankyrin repeats required for association with α and the carboxyl-terminal domain that dictates the ability to form homodimers. This region is common to all four NRF-2 β and γ subunits and is therefore consistent with the equal ability of all four to activate transcription when fused to GAL4.

These results contrast with those obtained with the E4TF1 using an in vitro transcription assay (26). In this system, E4TF1-53 stimulated transcription to a much greater degree than E4TF1-47 (NRF-2 γ_2 [Table 1]). Although the domain that we describe here was found to activate transcription in a construct lacking the homodimerization domain, in other constructs the major activity appeared to reside within the homodimerization domain of E4TF1-53. These workers concluded that the homodimerization and activation domains are coincident and that dimerization is required for transcriptional activation. By contrast, we observe no requirement for the homodimerization domain with use of GAL4 fusions in transfected cells. It had no activity on its own, and deletions removing it actually increased the ability of the GAL4 fusion protein to transactivate the reporter.

One explanation for the discrepancy is that in the in vitro transcription system, the amount of transcription obtained depends on the ability of the subunits to form hetero- and homodimers as well as on the presence of a functional activation domain. Because multiple ETS-binding sites were present in the transcription template (26), the differences in transcriptional activity between E4TF1-53 and -47 may simply reflect the large difference in binding affinity between heteromeric complexes containing these subunits. Therefore, mutations that interfere with hetero- or homodimer formation would be expected to markedly reduce the levels of transcription obtained from these templates by affecting DNA binding. Since the GAL4 fusion proteins are expected to bind DNA as homodimers (3), their ability to activate transcription is independent of interactions between β or γ subunits themselves and their effects on the formation of DNA-protein complexes.

Finally, the NRF-2 activation domain identified here has features in common with those of the glutamine-rich class of activation domain found in Sp1 (14). Recent studies indicate that activation domains of a given class are not functionally identical and that clusters of hydrophobic residues rather than the predominant glutamines may be more important for activation (14). The minimal NRF-2 activation domain is only 15% glutamine, but these residues are evenly spaced and clustered with residues having bulky hydrophobic side chains such as isoleucine and valine. It has been suggested that such periodicity of hydrophobic regions mediates interactions with several targets within the transcriptional apparatus (5, 14, 30). Thus, the presence of these features and their resemblance to those found in other well-characterized transcription factors are fully consistent with the assignment of a transcriptional activation function to this region.

ACKNOWLEDGMENTS

We thank R. Morimoto for cDNA libraries and Stephanie Hsu and Ricardo Monzon for RNA samples.

This work was supported by Public Health Service grant GM32525-12 from the National Institutes of Health. R.C.S. is supported by Faculty Research Award FRA-361 from the American Cancer Society. S.G. is a predoctoral trainee on Carcinogenesis Training Program 5T32 CA09560 from the National Institutes of Health.

REFERENCES

1. **Andrews, N. C., and D. V. Faller.** 1991. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. **19:**2499.
2. **Bottinger, E. P., C. S. Shelley, O. C. Farokhzad, and M. A. Arnaout.** 1994. The human β2 integrin CD18 promoter consists of two inverted ets *cis* elements. Mol. Cell. Biol. **14:**2604–2615.
3. **Carey, M., H. Kakidani, J. Leatherwood, F. Mostashari, and M. Ptashne.** 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. J. Mol. Biol. **209:**423–432.
4. **Chau, C. A., M. J. Evans, and R. C. Scarpulla.** 1992. Nuclear respiratory factor 1 activation sites in genes encoding the gamma-subunit of ATP synthase, eukaryotic initiation factor 2α, and tyrosine aminotransferase. Specific interaction of purified NRF-1 with multiple target genes. J. Biol. Chem. **267:**6999–7006.
5. **Cress, W. D., and S. J. Triezenberg.** 1991. Critical structural elements of the VP16 transcriptional activation domain. Science **251:**87–90.
6. **De la Brousse, F. C., E. H. Birkenmeier, D. S. King, L. B. Rowe, and S. L. McKnight.** 1994. Molecular and genetic characterization of GABPβ. Genes Dev. **8:**1853–1865.
7. **Denhardt, D. T.** 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23:**641–646.
8. **Dupriez, V. J., M. I. Darville, I. V. Antoine, A. Gegonne, J. Ghysdael, and G. G. Rousseau.** 1993. Characterization of a hepatoma mRNA transcribed from a third promoter of a 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase-encoding gene and controlled by *ets* oncogene-related products. Proc. Natl. Acad. Sci. USA **90:**8224–8228.
9. **Evans, M. J., and R. C. Scarpulla.** 1988. Both upstream and intron sequence elements are required for elevated expression of the rat somatic cytochrome *c* gene in COS-1 cells. Mol. Cell. Biol. **8:**35–41.
10. **Evans, M. J., and R. C. Scarpulla.** 1989. Interaction of nuclear factors with multiple sites in the somatic cytochrome *c* promoter. Characterization of upstream NRF-1, ATF and intron Sp1 recognition sites. J. Biol. Chem. **264:**14361–14368.
11. **Evans, M. J., and R. C. Scarpulla.** 1990. NRF-1: a *trans*-activator of nuclear-encoded respiratory genes in animal cells. Genes Dev. **4:**1023–1034.
12. **Fisher, R. P., and D. A. Clayton.** 1988. Purification and characterization of human mitochondrial transcription factor 1. Mol. Cell. Biol. **8:**3496–3509.
13. **Genuario, R. R., D. E. Kelley, and R. P. Perry.** 1993. Comparative utilization of transcription factor GABP by the promoters of ribosomal protein genes rpL30 and rpL32. Gene Expression **3:**279–288.
14. **Gill, G., E. Pascal, Z. H. Tseng, and R. Tjian.** 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the $dTAF_{II}110$ component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91:**192–196.
15. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
16. **Karim, F. D., L. D. Urness, C. S. Thummel, M. J. Klemsz, S. R. McKercher, A. Celada, C. Van Beveren, R. A. Maki, C. V. Gunther, J. A. Nye, and B. J. Graves.** 1990. The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes Dev. **4:**1451–1453.
17. **LaMarco, K., C. C. Thompson, B. P. Byers, E. M. Walton, and S. L. McKnight.** 1991. Identification of Ets- and notch-related subunits in GA binding protein. Science **253:**789–792.
18. **Lillie, J. W., and M. R. Green.** 1989. Transcription activation by the adenovirus E1a protein. Nature (London) **338:**39–44.
19. **Macleod, K., D. Leprince, and D. Stehelin.** 1992. The *ets* gene family. Trends Biochem. Sci. **17:**251–256.
20. **Marchioni, M., S. Morabito, A. L. Salvati, E. Beccari, and F. Carnevali.** 1993. XrpFI, an amphibian transcription factor composed of multiple polypeptides immunologically related to the GA-binding protein α and β

subunits, is differentially expressed during *Xenopus laevis* development. Mol. Cell. Biol. **13:**6479–6489.

21. **Parisi, M. A., B. Xu, and D. A. Clayton.** 1993. A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both in vivo and in vitro. Mol. Cell. Biol. **13:**1951–1961.
22. **Pognonec, P., H. Kato, H. Sumimoto, M. Kretzschmar, and R. G. Roeder.** 1991. A quick procedure for purification of functional recombinant proteins over-expressed in *E. coli*. Nucleic Acids Res. **19:**6650.
23. **Sadasivan, E., M. M. Cedeno, and S. P. Rothenberg.** 1994. Characterization of the gene encoding a folate-binding protein expressed in human placenta. J. Biol. Chem. **269:**4725–4735.
24. **Sadowski, I., and M. Ptashne.** 1989. A vector for expressing GAL4(1–147) fusions in mammalian cells. Nucleic Acids Res. **17:**7539.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
26. **Sawada, J.-I., M. Goto, C. Sawa, H. Watanabe, and H. Handa.** 1994. Transcriptional activation through the tetrameric complex formation of E4TF1 subunits. EMBO J. **13:**1396–1402.
27. **Shadel, G. S., and D. A. Clayton.** 1993. Mitochondrial transcriptional initiation. Variation and conservation. J. Biol. Chem. **268:**16083–16086.
28. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
29. **Thompson, C. C., T. A. Brown, and S. L. McKnight.** 1991. Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. Science **253:**762–768.
30. **Tjian, R., and T. Maniatis.** 1994. Transcriptional activation: a complex puzzle with few easy pieces. Cell **77:**5–8.
31. **Virbasius, C. A., J. V. Virbasius, and R. C. Scarpulla.** 1993. NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators. Genes Dev. **7:**2431–2445.
32. **Virbasius, J. V., and R. C. Scarpulla.** 1991. Transcriptional activation through ETS domain binding sites in the cytochrome *c* oxidase subunit IV gene. Mol. Cell. Biol. **11:**5631–5638.
33. **Virbasius, J. V., and R. C. Scarpulla.** 1994. Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. Proc. Natl. Acad. Sci. USA **91:**1309–1313.
34. **Virbasius, J. V., C. A. Virbasius, and R. C. Scarpulla.** 1993. Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. Genes Dev. **7:**380–392.
35. **Wang, K., K. M. Bohren, and K. H. Gabbay.** 1993. Characterization of the human aldose reductase gene promoter. J. Biol. Chem. **268:**16052–16058.
36. **Wasylyk, B., S. L. Hahn, and A. Giovane.** 1993. The Ets family of transcription factors. Eur. J. Biochem. **211:**7–18.
37. **Watanabe, H., J.-I. Sawada, K.-I. Yano, K. Yamaguchi, M. Goto, and H. Handa.** 1993. cDNA cloning of transcription factor E4TF1 subunits with Ets and notch motifs. Mol. Cell. Biol. **13:**1385–1391.
38. **Watanabe, H., T. Wada, and H. Handa.** 1990. Transcription factor E4TF1 contains two subunits with different functions. EMBO J. **9:**841–847.