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

Joseph V. Virbasius, Ching-man A. Virbasius, and Richard C. Scarpulla

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

The ETS domain proteins are a diverse family of transcriptional activators that have been implicated recently in the expression of cell-specific and viral promoters. Nuclear respiratory factor 2 (NRF-2) is a nuclear transcription factor that activates the proximal promoter of the rat cytochrome c oxidase subunit IV (RCO4) gene through tandem sequence elements. These elements conform to the consensus for high-affinity ETS domain recognition sites. We have now purified NRF-2 to homogeneity from HeLa cells and find that it consists of five polypeptides, only one of which has intrinsic DNA-binding ability. The others participate in the formation of heteromeric complexes with distinct binding properties. NRF-2 also specifically recognizes multiple binding sites in the mouse cytochrome c oxidase subunit Vb (MCO5b) gene. As in the functionally related RCO4 gene, tandemly arranged NRF-2 sites are essential for the activity of the proximal MCO5b promoter, further substantiating a role for NRF-2 in respiratory chain expression. Determination of peptide sequences from the various subunits of HeLa NRF-2 reveals a high degree of sequence identity with mouse GA-binding protein (GABP), a multisubunit ETS domain activator of herpes simplex virus immediate early genes. A cellular role in the activation of nuclear genes specifying mitochondrial respiratory function is thus assigned to an ETS domain activator of viral promoters.

[Key Words: ETS domain; oxidative phosphorylation; nuclear respiratory factors; mitochondria; transcription]

Received December 1, 1992; revised version accepted January 4, 1993.

All but a handful of the hundreds of gene products required for mitochondrial oxidative function, as well as for mitochondrial replication, transcription, and translation, are of nuclear origin (Clayton 1991; Hatefi 1985). While investigating mechanisms contributing to the coordinate regulation of these nuclear genes, Evans and Scarpulla (1989) first identified nuclear respiratory factor 1 (NRF-1) as a transcriptional activator of the rat cytochrome c gene. Functional NRF-1 sites have since been found in nuclear genes encoding subunits of four of the five mitochondrial respiratory complexes and an essential RNA component of the mitochondrial replication machinery (Evans and Scarpulla 1990; Chau et al. 1992). More recently, analysis of the rat cytochrome c oxidase subunit IV gene (RCO4) revealed that a major element of its core promoter consists of tandem binding sites for a factor designated as nuclear respiratory factor 2 (NRF-2) (Virbasius and Scarpulla 1990, 1991). A functional NRF-2 site was also identified in the human ATP synthase β-subunit gene. Essential to each NRF-2 site is the GGAA motif characteristic of the recognition sequences for the family of ETS domain proteins. This family is defined by a conserved 85-amino-acid sequence similar to the DNA-binding domain of the ets-1 proto-oncogene product (Karim et al. 1990; Nye et al. 1992). NRF-2-binding sites were shown to form similar complexes and to cross-compete with the prototypical Ets-1-binding site in the Moloney murine sarcoma virus long terminal repeat (LTR) (Virbasius and Scarpulla 1991). Several additional ETS domain proteins, including a number of activators of tissue-specific genes and several that are involved in the expression of viral genomes, have been identified (for review see Macleod et al. 1992). Although most of these appear to bind as monomers, at least one, the GA-binding protein (GABP) (Thompson et al. 1991; LaMarco et al. 1991), is comprised of heterologous subunits that have different functional domains.

To investigate the possible relationship between NRF-2 and the other ETS domain activators, we have
purified the protein to homogeneity by taking advantage of its high affinity for specific sites within the RCO4 promoter region. Primary structure analysis of the NRF-2 subunits reveals a high degree of sequence identity with the mouse GABP subunits. Moreover, a closely spaced pair of NRF-2 sites in proximity to the transcription start sites of a second oxidase subunit gene, MCOSb, are strikingly similar to those found in the RCO4 gene and are essential for function of the MCOSb core promoter. These observations establish NRF-2 as the human homolog of mouse GABP and further substantiate its cellular function in the expression of respiratory chain subunits.

Results

Specific recognition and transcriptional activation of cytochrome oxidase promoter elements by purified NRF-2

We purified HeLa NRF-2 to apparent homogeneity using a series of chromatographic steps culminating in two serial passages through a site-specific DNA affinity column (Fig. 1A). The affinity matrix was prepared with a catenated RCO4 + 13/+36 oligonucleotide representing one of the tandem binding sites in RCO4. Five polypeptides, designated α [56 kD], β1 [48 kD], β2 [46 kD], γ1 [41 kD] and γ2 [39 kD], are reproducibly observed in eluates from each pass of the affinity column (lanes 6, 7). The mass of the 56-kD α subunit agrees well with that of the 55-kD protein predicted from UV cross-linking (Virbasius and Scarpulla 1991).

NRF-2 activity reveals itself as three distinct complexes in mobility retardation assays. The major complexes indicated as bands 2 and 3 in Figure 1B migrate as a closely spaced doublet. The band 1 complex often appears heterogeneous but, as shown below, is the product of a protein species of unique molecular mass. The same pattern of shifted complexes found in nuclear extracts is preserved throughout the purification scheme, and at each stage of the purification, all three complexes are competitively displaced by an excess of unlabeled RCO4 + 13/+36 oligomer [Fig. 1B, even-numbered lanes]. The DEAE-agarose fraction contains a nonspecific complex comigrating with band 1, but this is removed by further purification. Thus, the various complexes arise either from the binding of the individual NRF-2 polypeptides to their recognition site or from the association of polypep-
tide subunits to generate different multimeric forms of NRF-2.

Although NRF-2 was purified on the basis of its high affinity to RCO4 recognition sites, it was of interest to demonstrate that the pure protein could activate transcription through these same sites. Tandem NRF-2 sites from RCO4 were thus cloned into a cytochrome c expression vector deleted of its upstream promoter and tested for NRF-2-activated transcription in vitro. The results shown in Figure 1C demonstrate a significant [approximately seven-fold] NRF-2-dependent activation of transcription (lanes 7–10). The low level of stimulation (approximately two-fold) of the −66 control vector lacking the RCO4 NRF-2 sites (lanes 3–6) possibly results from weak binding to multiple GGAAG NRF-2 core motifs present in the bacterial vector sequences. Nevertheless, the in vitro transcriptional stimulation by NRF-2 acting on the RCO4 sites is significantly above this background level. The transcripts initiate at the precise position observed for the in vivo cytochrome c transcripts in liver RNA (lane 2) indicating that they accurately reflect the activation of the normal initiation complex. Thus, specific recognition of the RCO4 sites by purified NRF-2 coincides with its ability to stimulate transcription through the same recognition sites in vitro.

Multiple NRF-2–DNA complexes arise through heteromultimers of NRF-2 subunits

A clue to the nature of the multiple shifted complexes came from the distribution of β and γ polypeptides in individual fractions eluted from the affinity resin [Fig. 2]. Upon elution of NRF-2 from the final affinity column, the faster migrating band 2 complex is formed with the protein subunits present in earlier fractions of the salt gradient [Fig. 2, lower panel, fractions 10–16] compared with the slower migrating band 3 complex, which appears later in the gradient [fractions 16–26]. Although the α polypeptide is visible by SDS-PAGE in each active fraction [upper panel], γ1 and γ2 coelute with the band 2 complex while β1 and β2 coelute with the band 3 complex as detected by mobility retardation assay [lower panel]. Where a mixture of all five polypeptides is present in a single gradient fraction [e.g., see fraction 16] both band 2 and 3 complexes are detected. This observation suggests that the various NRF-2 complexes result from the specific association of distinct combinations of subunits.

To investigate the association of the purified subunits with the different shifted complexes, affinity-purified NRF-2 was fractionated on an SDS gel, and the individual protein subunits were eluted from gel slices and renatured. Activity of the renatured proteins was assayed by mobility retardation with the radiolabeled RCO4 + 13/+36 DNA-DNA binding site (Fig. 3). Despite the fact that the α-subunit was excised from the gel as a single band of 56 kD, the renatured polypeptide gave rise to the multiple complexes [lane 2], which comigrate with the similarly heterogeneous band 1 complex present in native NRF-2 fractions [lane 1]. When added individually, neither β-nor γ-subunits forms a complex with the DNA-binding site [lanes 3–6] nor do pairwise combinations of β and γ polypeptides result in DNA binding [not shown]. When each individual β and γ polypeptide is added to the renatured α polypeptide, however, a complex is generated that comigrates with either band 3 [lanes 7,8] or band 2 [lanes 9,10] of the native fraction [lane 1]. As predicted from the assays of the gradient fractions [Fig. 2], the β polypeptides participate in the slower migrating band 3 complex, whereas the α/γ mixture produces the band 2 complex. Each of the complexes is specific, as indicated by competition with excess NRF-2-binding site (not shown). These results establish that the predominant forms of the native NRF-2 factor consist of at least two polypeptide subunits, only one of which is required for DNA binding.

The β-subunits of heteromeric NRF-2 complexes enhance binding to tandem recognition sites

Mutational analysis of the RCO4 promoter indicated cooperative NRF-2 binding to tandem sites [Virbasius and Scarpulla 1991]. To investigate whether purified NRF-2 displays a similar enhanced binding to tandem sites, DNase I footprinting was performed by use of varying amounts of the pure native protein. The footprint pat-
NRF-2 activation of cytochrome oxidase genes

Figure 3. Reconstitution of NRF-2 complexes with purified renatured subunits. (Lane 1) The pattern of shifted complexes obtained with second-pass affinity fraction [Aff. PII] and radiolabeled RCO4 + 36 oligonucleotide; (Lanes 2–10) the pattern of shifted complexes obtained with the purified NRF-2 polypeptides that were renatured following excision from a preparative denaturing gel. The renatured subunits present in the binding reaction, as indicated above each lane, were assayed using radiolabeled RCO4 + 36 oligonucleotide. The positions of shifted complexes detected by electrophoresis on a native acrylamide gel are designated as bands 1, 2, and 3 at right.

The pattern was compared between an RCO4 promoter fragment with intact NRF-2 sites designated A and B and the same fragment carrying site-directed point mutations in the B site (Fig. 4). With the intact RCO4 fragment (lanes 1–6), nearly complete protection of both the A (5') and B (3') NRF-2 sites is achieved by the addition of only 1 µl (~2.5 ng) of affinity-purified protein. In contrast, even partial protection of the intact A site in the MUTB fragment requires much more of the purified protein and is still incomplete when 20-fold more affinity fraction is added to the binding reaction. Thus, the purified protein exhibits a marked preference for tandem NRF-2 recognition sites.

Given the enhanced DNase I footprinting activity with the native factor on tandem sites, it was of interest to determine whether the various reconstituted forms of NRF-2 were equally proficient in binding to a promoter fragment containing intact A and B sites. Accordingly, renatured α alone or α reconstituted with equal activities (calibrated by measuring the ability to form band 2 or band 3 complexes in mobility retardation) of each of the β- and γ-subunits was added to a labeled RCO4 fragment with binding sites A and B intact (Fig. 5). When added alone, the α-subunit did not protect the NRF-2-binding sites from DNase I cleavage under these conditions (lane 2). When either β₁ or β₂ was added in addition to the same amount of α-subunit, however, complete protection of both binding sites was observed (lanes 3, 4). Although the γ-subunits were capable of reconstituting the band 2 complex with α as shown in Figure 3, neither was able to stimulate the interaction of α with tandem binding sites as measured by DNase I footprinting (lanes 5, 6). Furthermore, no footprint was obtained using any combination of subunits with the MUTB fragment having only one NRF-2 site (data not shown). We conclude that the β-subunits of NRF-2 form a complex with the α-subunit and enhance binding when tandem sites are available. With α alone or the α/γ complexes, tandem sites appear to be recognized independently. The increased amount of NRF-2 binding activity required to obtain a partial footprint on a single site (Fig. 4) accounts for the inability to obtain protection using renatured subunits on a single or functionally independent tandem sites because the renatured polypeptides recover only a fraction of the original activity. These results suggest that the function of the β- and γ-subunits is to modulate the DNA binding of α to tandem recognition sites within promoters.

Figure 4. Preferential binding of purified NRF-2 to tandem recognition sites in the RCO4 promoter region. End-labeled fragments from an RCO4 promoter having both NRF-2-binding sites intact [RCO4] or containing site-directed point mutations in the GGAA motif of the B-binding site [RCO4MUTB] were subjected to DNase I footprinting. Binding reactions containing varying amounts of second-pass affinity-purified NRF-2 [2.5 ng/µl] as indicated above each lane, were treated with DNase I, and the resulting fragments were separated on denaturing gels. The A (5') and B (3') NRF-2-binding sites of the RCO4 promoter are indicated at right.
Virbasius et al.

Figure 5. The β-subunits of NRF-2 are required for its enhanced binding to tandem recognition sites in the RCO4 promoter region. DNase I footprinting was carried out by use of the RCO4 promoter fragment with both NRF-2 sites intact. End-labeled RCO4 fragment was incubated with renatured α-subunit alone (lane 2) or with the same amount of α-subunit plus each of the smaller NRF-2 β- and γ-subunits, as indicated above lanes 3–6. Positions of the A and B NRF-2-binding sites are indicated at right.

Interaction of purified NRF-2 with a tandem array of binding sites in proximity to the start of cytochrome c oxidase subunit Vb transcription

Examination of a recently published sequence of the mouse cytochrome c oxidase subunit Vb (MCO5b) promoter (Basu and Avadhani 1991) revealed the presence of four potential NRF-2-binding sites within 60 nucleotides of the sites of transcription initiation (Fig. 6). To test the ability of these sequences to interact specifically with purified NRF-2 in vitro, synthetic oligomers containing each were used as competitors for binding of affinity-purified NRF-2 to an RCO4 NRF-2-binding site. As shown in Figure 7, lanes 1–6 each of the MCO5b sites was effective in preventing formation of the NRF-2/DNA complex when the radiolabeled RCO4 +13/+36 oligomer was used as a probe. The observed competition was specific because under the same conditions, a control oligomer, hCC1 −181/−204, which is identical to the RCO4 binding site except for a single G → C change in the essential GGAA core sequence (Virbasius and Scarpulla 1991), was unable to compete. Furthermore, labeled oligomers from each of the MCO5b NRF-2 sites yielded a pattern of shifted complexes (lanes 7,11,15) similar to that formed with the RCO4-binding site (lane 1). Each of these complexes is effectively competed by a 200-fold molar excess of the same unlabeled oligomer (lanes 8,12,16) and with the RCO4-binding site (lanes 9,13,17) but not with the negative control hCC1 oligomer (lanes 10,14,18). The MCO5b −59/−39 oligomer differs from the others in that it has two closely spaced NRF-2 sites, which are designated as MCO5b sites A and B in Figure 6. This oligomer forms additional low-mobility complexes visible as a closely spaced doublet above bands 2 and 3 in lanes 7 and 10, which are also specific. These likely result from binding of NRF-2 to both of the closely spaced GGAA-binding sites in this sequence (see below).

To verify that purified NRF-2 binds the newly identified MCO5b sites through guanine nucleotides in the GGAA motif, we carried out methylation interference assays (Fig. 8). With the RCO4 +13/+36 probe both the band 1 complex, which results from binding of free α-subunit, and the band 2 + 3 doublet, which corresponds to the multimeric complexes formed with α plus the β- or γ-subunit, displayed the same guanine contacts as observed previously for RCO4 using a crude NRF-2 preparation (Virbasius and Scarpulla 1991). The same pattern of contacts was also observed for the band 2 + 3 doublet derived from each of the MCO5b sites (Fig. 8; MCO5b +13/+36, MCO5b −22/−1, and MCO5b +13/+33), indicating that recognition by the purified

| Gene       | NRF-2 recognition sequence |
|------------|---------------------------|
| RCO4 site A | +11 A C C G G A A A G G +2 |
| RCO4 site B | +29 A C C G G A A A G G +20 |
| MCO5b site A | −57 G C A G G A A C G G −48 |
| MCO5b site B | −52 A G C G G A A A C G C −43 |
| MCO5b site C | −15 A A C G G A A T C G −6 |
| MCO5b site D | +19 C C G G A A A G T C G +28 |
| β-ATPsyn    | +599 A C C G G A A A G T T +590 |

Figure 6. Comparison of NRF-2-binding sites and their tandem arrangement in RCO4 and MCO5b proximal promoters. NRF-2 sites from the RCO4 and human β-ATP synthase genes are compared with those in the MCO5b promoter. Nucleotides identical to the RCO4-binding sites are in the shaded boxes. The tandem arrangement of NRF-2 sites (open lettered boxes) relative to the approximate positions of transcription initiation sites (arrows) for both MCO5b (Basu and Avadhani 1991) and RCO4 (Virbasius and Scarpulla 1990) genes is summarized below.
activation of cytochrome oxidase genes

Figure 7. Specificity of binding of purified NRF-2 to individual recognition sites present in both RCO4 and MCO5b promoter regions. Labeled oligonucleotides containing the RCO4- or MCO5b-binding sites, as indicated above each series of lanes, were incubated with affinity-purified NRF-2, and the resulting complexes were separated on native polyacrylamide gels. Binding reactions contained either no competitor (lanes 1, 7, 11, 15) or a 100-fold molar excess of the unlabeled competitor oligonucleotides, as indicated above each individual lane. The positions of the three NRF-2 shifted complexes are indicated as bands 1–5 at right.

protein is the same in each case. Of interest is the pattern obtained with MCO5b -59/-39, which contains closely spaced tandem NRF-2 sites [MCO5b sites A and B of Fig. 6]. In addition to the band 2 + 3 doublet, bound DNA was purified from the slower migrating doublet (designated as bands 4 and 5 in Fig. 7), which likely results from occupancy of both GGAA motifs by NRF-2. In both cases, bound DNA exhibits less severe interference at each of the guanine nucleotides in the two GGAA motifs and the G between them. For the band 2 + 3 doublet, incomplete interference likely results from binding of NRF-2 to only one of the two GGAA motifs [in which case methylation at the other motif would still be present in the bound DNA pool]. For the band 4 + 5 doublet, interference is greater still, but not complete, possibly reflecting the dispensability of any one contact when two NRF-2 molecules are bound to adjacent sites. Therefore, the multiple NRF-2 sites found in the MCO5b promoter region are authentic targets for NRF-2 recognition in vitro.

Activation of the MCO5b promoter by multiple NRF-2 recognition sites

To investigate the potential function of the MCO5b NRF-2 sites, oligomers of each were cloned upstream of a truncated rat cytochrome c promoter linked to a CAT expression cassette and transfected into COS cells. As shown in Table 1, the RCO4 +13/+36 NRF-2 oligomer gives about four-fold stimulation over the promoter with no binding site added or the negative control hCC1 -181/-204 oligomer, which does not bind NRF-2 in vitro. Under the same conditions, all of the MCO5b oligomers gave significant levels of promoter stimulation when present as single sites in the heterologous promoter. The ability of the MCO5b NRF-2 sites to function independently in these transfection assays is consistent with the in vitro-binding experiments, suggesting that they are functional components of the MCO5b promoter.

Given the remarkable similarity in the proximal promoters of RCO4 and MCO5b genes [see Fig. 6] it was of interest to define further the contribution of the various MCO5b NRF-2 sites within the proper promoter context. To that end, the proximal MCO5b promoter region between -59 and +33 was assembled in modular fashion from overlapping synthetic oligomers and ligated to a rat cytochrome c/chloramphenicol acetyltransferase (CAT) expression vector. This vector was truncated at its 5' end to position +17, thus replacing all cytochrome c promoter sequences and start sites with those resident in the MCO5b -59/+33 sequence [Fig. 10, below]. Mutant versions of each module encompassing the various MCO5b NRF-2 sites, A, B, C, and D, were synthesized by replacement of the GGAA motif with TTAA and the mutant modules substituted both individually and in combinations of two or three in the expression vector. Because of the close proximity of the A and B sites, they were incorporated into the same module with both GGAA motifs disrupted in the mutated form.

DNase I footprinting was carried out to determine whether these promoter constructions bind affinity-pu-
Virbasius et al.

Figure 8. Methylation interference analysis of NRF-2 binding to RCO4 and MCO5b recognition sites. End-labeled DNA fragments containing the cloned RCO4 +13/+36, MCO5b -59/-39, MCO5b -22/-1, and MCO5b +13/+33 oligonucleotides were partially methylated and subjected to preparative scale mobility retardation analysis using purified NRF-2. Both free DNA (F) and DNA isolated from bound complexes (B) were cleaved with piperidine, and the resulting fragments were separated on sequencing gels. Bands 1-5 refer to distinct NRF-2/DNA complexes resolved by mobility retardation assays as designated in Fig. 7. • Guanosine bases that, when methylated, strongly inhibited NRF-2 binding. The DNA sequence of each site, along with the positions of guanine nucleotide contacts, is summarized below.

Figure 8. Methylation interference analysis of NRF-2 binding to RCO4 and MCO5b recognition sites. End-labeled DNA fragments containing the cloned RCO4 +13/+36, MCO5b -59/-39, MCO5b -22/-1, and MCO5b +13/+33 oligonucleotides were partially methylated and subjected to preparative scale mobility retardation analysis using purified NRF-2. Both free DNA (F) and DNA isolated from bound complexes (B) were cleaved with piperidine, and the resulting fragments were separated on sequencing gels. Bands 1-5 refer to distinct NRF-2/DNA complexes resolved by mobility retardation assays as designated in Fig. 7. • Guanosine bases that, when methylated, strongly inhibited NRF-2 binding. The DNA sequence of each site, along with the positions of guanine nucleotide contacts, is summarized below.

RCO4 +13/+36 ACCGGAAGAG
MCO5b -59/-39 GGAAGAGCAGGC
MCO5b -22/-1 AAGGTGATCG
MCO5b +13/+33 CCGGAAATGC

Stimulation of a truncated cytochrome c promoter by cloned oligonucleotides

| Oligonucleotide          | Relative activity * |
|--------------------------|---------------------|
| None                     | 1.0                 |
| hCC1 -181/-204           | 1.1 ± 0.2           |
| RCO4 +13/+36             | 4.4 ± 1.5           |
| MCO5b -59/-39 site A/B   | 14.3 ± 7.4          |
| MCO5b -22/-1 site C      | 2.8 ± 1.2           |
| MCO5b +13/+33 site D     | 23.3 ± 7.7          |

*Normalized CAT activity relative to the construct with no oligonucleotide [pRC4CATB/-668A] expressed as the average of at least four determinations ± s.d.
Figure 9. Recognition of specific sites in the MCOSb proximal promoter by purified NRF-2. DNase I footprinting was performed on MCOSb promoter fragments that were synthesized in modular fashion and cloned into expression vectors. End-labeled fragments containing the intact MCOSb 59/-33 sequence (WT) or mutated versions containing substitutions within the GGAA motif of each of the NRF-2-binding sites (MUT A/B, MUT C, MUT D) were incubated either with no added factor (- lanes) or with affinity-purified NRF-2 before treatment with DNase I (+ lanes). Positions of the A/B, C and D NRF-2 recognition sites are indicated on right.

Neither of these homologous regions is within the conserved ETS domain region of GABPα, making it unlikely that we fortuitously chose conserved binding domain sequences in an otherwise unrelated protein.

Given the close similarity of the peptide maps of the β and γ preparations, we analyzed a peak common to both and several unique to each map. One common peak, βγ80, contained a 21-residue peptide, which is identical to the common sequence of the GABP β1- and β2-subunits (residues 105–125). A peptide from β30, a peak unique to NRF-2 β, contained a perfect match to 8 residues (367–374) of the carboxy-terminal region unique to GABP β1. Two peaks unique to NRF-2 γ were also analyzed. The first, γ28, contained a 9-residue peptide, which except for a single Cys→Ser change, matches residues 327–335 of the GABP β2 sequence. However, the divergence between GABP β1 and β2 follows residue 333, and thus only one difference (Glu or Val at position 334) is included within the span matched by the NRF-2 γ28 peptide. We therefore determined the sequence of an additional peptide from a peak unique to NRF-2-γγ. This peptide, γ60, matches 6 of 7 residues of GABP β2 sequence 336–342 and is therefore similar to sequences entirely unique to GABP β2. Thus, it would appear that at least one of the differences between the NRF-2 β- and γ-subunits is analogous to the sequence difference between the GABP β1- and β2-subunits. The precise assignment of homologs among the small subunits, however, awaits further molecular genetic and biochemical analysis [see Discussion]. Nevertheless, these results indicate that NRF-2 is the human homolog of mouse GABP.

Discussion

NRF-2 as a nuclear activator of respiratory chain expression

Analysis of the promoters of cytochrome c and cytochrome oxidase subunit IV genes [Evans and Scarpulla 1988, 1989; Virbasius and Scarpulla 1991] has led to the identification of the nuclear respiratory factors NRF-1 and NRF-2. In addition to the cytochrome c gene, functional NRF-1 recognition sites have been found in genes encoding subunits of respiratory complexes III, IV, and V, and the RNA subunit of MRP endonuclease, which participates in mitochondrial DNA replication [Evans and Scarpulla 1990; Chau et al. 1992]. We also find just upstream of the NRF-2 sites in the MCOSb gene a perfect match to an NRF-1 consensus-binding site (−92 CGCA-CATGCAGCA−103), which is functional in both NRF-1-binding and transcriptional activation [J. Virbasius, unpubl.]. Thus, the MCOSb gene represents the first respiratory chain gene whose expression appears to be highly dependent on both NRF-1 and NRF-2 recognition sites.

The tandem arrangement of NRF-2 sites in MCOSb is similar to those characterized previously in the RCO4 gene [Virbasius and Scarpulla 1991], the only other vertebrate cytochrome oxidase gene whose promoter region has been investigated. A minimal promoter analyzed for both genes (59 and 77 bp upstream of the first transcrip-
Figure 10. Promoter activities of MCO5b constructs carrying intact or mutant NRF-2-binding sites. The sequences from -59 to +33 of the MCO5b proximal promoter were synthesized with either intact NRF-2-binding sites or with sites mutated in the GGAA core sequence to TTAA. The thymine nucleotide substitutions in each mutated promoter sequence are underlined. Intact and mutated versions of the promoter were fused to the +17 position of the first exon of a rat cytochrome c/CAT expression vector pRC4CAT/+17 and subsequently transfected into COS cells. The resulting CAT activities were measured and normalized to the amount of CAT plasmid DNA recovered in the Hirt supernatants to correct for differences in transfection efficiency. Results are given as the average ± S.D. of three to five determinations. The first transcription start site is designated as +1.

Figure 11. Peptide mapping and microsequencing of NRF-2 subunits. (A) Products of in situ tryptic digestion of the NRF-2 α-polypeptide eluted from a reverse-phase HPLC column. The arrow marks the peak from which two α-subunit peptide sequences were derived. (B) Superimposed peptide maps of NRF-2 β and γ-subunits. Peptide sequencing was performed on material from the numbered peaks indicated by the arrows. (C) Sequences of NRF-2 subunits. Sequences of peptides from the peaks indicated in the maps above are compared with the known GABP subunit sequences [LaMarco et al. 1991]. Residues in NRF-2 that differ from GABP are underlined.
Assignment of NRF-2 as an ETS domain activator and the human homolog of GABP

The identification of NRF-2 as an activator of cytochrome oxidase expression defines a novel regulatory function for ETS domain proteins. Several such proteins have now been cloned and, in some cases the recognition sites for the recombinant proteins have been defined extensively (for review, see Macleod et al. 1992). Because of the high degree of cross-recognition among the ETS domain proteins for similar binding sites, it is important to identify the major molecular species responsible for NRF-2 activity. Despite the possibility of cross-recognition with other ubiquitous ETS-proteins, our affinity-purified fractions contain in detectable quantities only the five NRF-2 polypeptides. Their assignment as NRF-2 subunits is based on the ability to account for all of the NRF-2 complexes detected in crude nuclear extracts by reconstituting each complex from renatured polypeptides that were excised from denaturing gels. We have identified the primary sequences of these polypeptides as the human homologs of the subunits of mouse GABP. Although this does not exclude the possibility of other ETS domain proteins interacting with NRF-2 sites in vivo, it argues that the strongest and/or most abundant activity can be ascribed to this factor. In addition, the ability of NRF-2 to activate transcription through the RCO4 sites in vitro is also fully consistent with a biological role in cytochrome oxidase expression.

It should be noted that discrepancies regarding the precise relationship between the small GABP (β1 and β2) and NRF-2 (β1, β2, γ1, and γ2) subunits still remain. Preparations of GABP from rat liver were reported to contain three polypeptides between 43 and 68 kD, which are similar in apparent molecular mass to the α, β1, and β2 polypeptides of NRF-2 [LaMarco and McKnight 1989]. The bacterial products of 41 and 37 kD expressed from the full-length GABP β1 and β2 cDNAs [Thompson et al. 1991], however, correspond most closely to the NRF-2 β- and γ-subunits, respectively. This is consistent with our finding of peptide differences between NRF-2 β and γ analogues to those predicted from the GABP β1 and β2 sequences [LaMarco et al. 1991]. Furthermore, functional differences between the NRF-2 α/β and α/γ complexes correspond to properties expected for the GABP α/β and α/β2 complexes. We presume that the difference in migration between NRF-2 β1 and β2 or between γ1 and γ2 results from either a sequence difference that we did not discover at the protein level or perhaps some form of protein processing or modification.

Role of subunits in the binding properties of purified NRF-2

The NRF-2 β-subunits confer on the heteromeric complex the ability to bind avidly to tandem binding sites. A similar property was observed for the GABP β1 protein, which is required for avid binding to the GGAA repeats spaced at the herpes simplex virus (HSV) ICP4 gene [Thompson et al. 1991]. This property was attributed to the unique ability of the GABP β1 to form dimers in solution and to potentiate the formation of heterotetramers including two α-subunits. Such a complex was proposed to bind in a concerted manner to repeated recognition sequences. The tandem sites of the RCO4 promoter are arranged differently from those in the HSV gene. Measured from the center of each GGAA sequence, the two sites in RCO4 are 20 bp apart, or 1.9 turns of the DNA helix. Interestingly, we have observed that the same cooperative binding dependent on the β-subunit of NRF-2 applies to the C and D sites of the MCO5b promoter, which are spaced 30 bp, or 2.9 turns, apart [J. Virbasius, unpubl.]. Thus, the interaction of multiple binding complexes dependent on the β-subunits of NRF-2 or the analogous βγ-subunit of GABP must be quite flexible to permit concerted binding to closely spaced or separated sites. In the latter case, the two elements that display this property are spaced by an approximately integral number of helical turns, suggesting the necessity of sites on the same face of the helix. This dependence on integral spacing for cooperative binding has been observed in the classic case of the λ repressor [Hochschild 1992].

What, then, is the function of the alternative γ-subunits of NRF-2? As Thompson and co-workers pointed out, the relative expression of the subunits might significantly affect the function of the resulting complexes in cells [Thompson et al. 1991]. We have observed differences in the relative amounts of gel shift bands corresponding to the α/β and α/γ complexes in the various cell and tissue nuclear extracts [Virbasius and Scarpulla 1991]. Likewise, differences in the relative amounts of mRNAs provisionally assigned to GABP β1 and β2 are apparent in various tissues. Because the expression of most nucleus-encoded respiratory genes including cytochrome c oxidase subunits varies widely among tissues [Virbasius and Scarpulla 1990], it is tempting to speculate that the relative amounts of the various NRF-2 subunits contribute to this tissue-specific modulation.

Materials and methods

Purification and amino acid sequencing of NRF-2

Nuclear extracts were prepared [Dignam et al. 1983] from HeLa cells grown in spinner cultures to a density of 10⁶ cells/ml. One unit is the amount of NRF-2 activity required to shift 5 fmols of labeled RCO4 + 13/ + 36 oligonucleotide under standard mobility retardation conditions described below. Typically 900 mg of nuclear extract [4.6 mg protein/ml, sp. act., 0.020 U/μg] from 2 x 10¹¹ cells was diluted two fold in Tris-D (20 mm Tris-HCl, at pH 7.9, 20% glycerol, 0.2 mm EDTA, 0.5 mm phenylmethylsulfonyl fluoride, 0.5 mm dithiothreitol) and was applied to a DEAE–agarose (Bio-Rad) column (2.5 x 12 cm) equilibrated in Tris-D, 50 mm KCl. The column was washed with five bed volumes of the same buffer and protein eluted with a linear
gradient [five column volumes] of 50 mM to 1M KC1 in Tris-D. Fractions were assayed by mobility retardation analysis as described below, and fractions containing NRF-2 activity were pooled [final concentration, 2.0 mg protein/ml; sp. act., 0.033 U/µg], diluted to 100 mM KC1 with HEPES-D (same as Tris-D) except 20 mM HEPES–KOH, at pH 7.9 instead of Tris) and applied to a heparin–agarose [Sigma] column (2.5 × 25 cm) equilibrated in HEPES-D, 100 mM KC1. After washing with five bed volumes of the same buffer, NRF-2 was eluted with two to three volumes of HEPES-D, 250 mM KC1. This fraction (0.37 mg of protein/ml; sp. act., 0.061 U/µg) was adjusted to 100 mM KC1, and 0.1% NP-40, and applied to a DNA cellulose [Sigma] column (1.6 × 13.5 cm) equilibrated in HEPES-D, 100 mM KC1, and 0.1% NP-40. After washing with the same buffer, proteins were eluted with a 100 mM to 1 M gradient of KC1 in HEPES-D, 0.1% NP-40. Fractions containing NRF-2 activity were combined (0.61 mg of protein/ml; sp. act., 0.36 U/µg) and subjected to affinity chromatography.

A specific DNA affinity column (3-ml bed volume) was prepared from the RCO4 + 13/- 36 oligonucleotides as described [Kadonaga and Tjian 1986]. DNA–cellulose fractions containing NRF-2 were adjusted to 100 mM KC1, and 400 µg of a 2:1 mixture of native and denatured sonicated calf thymus DNA was added for each milliliter of DNA–cellulose fraction. This mixture was added for each milliliter of DNA–cellulose fraction. This mixture was reapplied to the affinity column equilibrated in HEPES-D, 100 mM KC1, and 0.1% NP-40. After washing with five bed volumes of the same buffer, proteins were eluted in a 100 mM to 1 M KC1 gradient in HEPES-D, 0.1% NP-40. Active fractions from this first pass were diluted to 100 mM KC1 and 100 µg of mixed native and denatured calf thymus DNA and 200 µg insulin were added for each milliliter of pooled fraction. This mixture was reapplied to the same affinity column, and NRF-2 was eluted as described for the first pass. Fractions containing NRF-2 activity were pooled [final concentration: 2.5 ng of protein/µl; sp. act., 100 U/µg] and stored in aliquots at −80 °C.

Affinity-purified NRF-2 was prepared for tryptic digestion and sequencing as described [Aebersold et al. 1997]. Subsequent digestion, peptide separation, and microsequencing was carried out by William S. Lane [Harvard Microchemistry Facility].

In vitro transcription

Nuclear extracts prepared as described above were concentrated by the addition of 0.33 g/ml of ammonium sulfate and recovery of the precipitated proteins by centrifugation [10,000g, 20 min]. The protein pellet was resuspended in 1/60 the original extract volume of Tris-D with 100 mM KC1 and dialyzed against two changes of 1000 volumes of the same buffer. In vitro transcription reactions were carried out as described [Shapiro et al. 1988] except that no labeled nucleotide was included, and the concentration of unlabeled UTP was increased to 100 µM. NRF-2 affinity column fractions were also dialyzed against Tris-D with 100 mM KC1 before use in transcription reactions. Reaction products were recovered by phenol–chloroform extraction and ethanol precipitation. Transcripts were assayed by primer extension [Virbasius and Scarpulla 1991] using a labeled oligonucleotide complementary to the last 18 nucleotides of the rat cytchrome c gene first exon.

Renaturation of NRF-2 polypeptides

Affinity-purified NRF-2 preparations were fractionated on a 10% acrylamide–SDS gel, and protein bands were visualized using the Fast Stain kit [Zoion Research]. Activity was recovered by a modification of the method of Hager and Burgess [1980] as described previously [Chau et al. 1992], except that renaturation was carried out by dialysis against two changes of 200 volumes of denaturation buffer without guanidine-HCl. NRF-2 complexes were reconstituted by direct mixing in the DNA-binding reaction.

Mobility retardation assays, DNase I, and methylation interference footprinting

Synthetic DNA oligomers used for DNA binding analyses were as follows:

| Oligonucleotide | Sequence                        |
|----------------|--------------------------------|
| RCO4 (+13/-36) | GATCCGGGACCCCGCCCTTCCGGGCGAA |
| hCcc (-181/-201) | GCCCTCTGGGGAAGAAGGCCCAGTTTTCGA |
| MCO5b (-59/-39) | GATCCTCGGGAGAGGGCCAGCTGCA |
| MCO5b (+22/-1) | GCCGGTCTTCGGCCCTTCTGCTGCA |
| MCO5b (+13/+33) | GATCTTGTTCCGGGCTGGTTTCCGCA |

Binding reactions were performed in TM buffer (25 mM Tris-HCl at pH 7.9, 6.25 mM MgCl2, 0.5 mM EDTA, 0.5 mM dithiothreitol, 50 mM KC1, 10% glycerol) for 15–30 min at room temperature. Reactions using affinity-purified or renatured proteins contained 5 µg/ml of sonicated calf thymus DNA and 50 µg/ml of BSA. For mobility retardation assays, 10 fmole of the above 32P-end-labeled oligonucleotides were incubated with protein in the above binding mixture. Competition binding reactions were performed by the addition of 2.0 pmole of unlabeled oligonucleotide before the addition of labeled oligonucleotide. The DNA–protein complexes were visualized by autoradiography after gel electrophoresis as described [Virbasius and Scarpulla 1991; Chau et al. 1992].

Methylation interference was performed on the above oligonucleotides that had been cloned in the polylinker of pGEM-7Zf (+) and excised by XhoI and MluI cleavage as described previously [Evans and Scarpulla 1990, Virbasius and Scarpulla 1991]. DNase I footprint-binding reactions were carried out by use of 5–10 fmoles of single-end-labeled DNA fragments from the RCO4 (226-bp EcoRI–NcoI) or MCO5b (345-bp BssHII–Asel) promoters in Z buffer (12.5 mM HEPES–KOH, at pH 7.9, 6.25 mM MgCl2, 0.5 mM dithiothreitol, 10% glycerol, 0.05% NP-40) with 2% polyvinyl alcohol. Following a 15- to 30-min binding reaction on ice, the sample was adjusted to 5 mM MgCl2, 2.5 mM CaCl2 and treated with 12.5–25 ng/ml of DNase I for 1 min. Cleaved DNA was extracted with phenol/chloroform/isoamyl alcohol, ethanol precipitated, and analyzed on 8% or 10% polyacrylamide–urea sequencing gels.

Promoter constructs and expression vectors

The vector pRC4CATBΔ/-66Δ [Evans and Scarpulla 1989] was used to measure stimulation of the truncated rat cytochrome c (RC4) promoter by the individual NRF-2-binding sites as described previously [Evans and Scarpulla 1990, Virbasius and Scarpulla 1991; Chau et al. 1992]. NRF-2 oligonucleotides from RCO4, MCO5b, and hCcc, cloned in pGEM-7Zf (+) were excised by Clal and Asp718I cleavage and ligated into pRC4CATBΔ/-66Δ that had been cut with the same enzymes. This positioned the NRF-2-binding sites 95 nucleotides upstream from the RC4 transcription start site. The promoterless expression vector pRC4CAT7/ + 17 was constructed by conversion of the AatII site in the first exon of expression vector pRC4CATBΔ/-215, +82/+100 [Evans and Scarpulla 1989] with an Asp718I linker and insertion of the fragment from this site to the Sphi site downstream of the rat
cytochrome c polyadenylation signals into the polylinker of the vector pGEM 7Zf (+) [Promega]. The MCO5b sequence from cytochrome c polyadenylation signals into the polylinker of pRC4CAT / +17, thus fusing the +33 first nucleotide kinase. Following ligation with T4 DNA ligase, the gel. This fragment was ligated to the exon. For constructions incorporating mutated NRF-2-binding sites, the following oligonucleotides were substituted for the lower strand of segment 4, were phosphorylated with T4 polynucleotide kinase. Following ligation with T4 DNA ligase, the gel. This fragment was ligated to the exon.

All of these, excluding the upper strand of segment 1 and the lower strand of segment 4, were phosphorylated with T4 polynucleotide kinase. Following ligation with T4 DNA ligase, the resulting 98-bp fragment was isolated from a polyacrylamide gel. This fragment was ligated to the BamH1 and Asp718I sites in the polylinker of pRC4CAT / +17, thus fusing the +33 first exon position of MCO5b to position +17 of the RC4CAT first exon. For constructions incorporating mutated NRF-2-binding sites, the following oligonucleotides were substituted for the above:

M1: GATCTGGGCGAGAAGCGCAAG
M2: GGGCGGAGAGAGCGGAAGA
M3: TCCGACAGCCTTGAAGTGCT
M4: TCTGCTACAGGCTTGAAGTGCT

Transfections

Growth of COS-1 cells, transfection by the calcium phosphate method [Graham and van der Eb 1973], and CAT assays were carried out as described previously [Evans and Scarpulla 1988, 1989]. Low-molecular-weight DNA was prepared from one-half of the transfected cells by the method of Hirt [1967]. CAT activities reported have been normalized to the amount of CAT expression efficiency as described previously [Evans and Scarpulla 1989].

Acknowledgments

We thank William S. Lane of the Harvard Microchemistry Facility for his invaluable guidance and expertise in performing the peptide mapping and amino acid sequencing of NRF-2 subunits. This work was supported by U.S. Public Health Service Grant GM32525-10 from the National Institutes of Health. R.C.S. is the recipient of Faculty Research Award FRA-361 from the American Cancer Society.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Aebersold, R.H., J. Leavitt, R.A. Saavedra, L.E. Hood, and S.B.H. Kent. 1987. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose. Proc. Natl. Acad. Sci. 84: 6970–6974.

Basu, A. and N.G. Avadhani. 1991. Structural organization of nuclear gene for subunit Vb of mouse mitochondrial cytochrome c oxidase. J. Biol. Chem. 266: 15450–15456.

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 20alpha, and tyrosine aminotransferase. Specific interaction of purified NRF-1 with multiple target genes. J. Biol. Chem. 267: 6999–7006.

Clayton, D.A. 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. 7: 453–478.

Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11: 1475–1489.

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.

——. 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.

——. 1990. NRF-1: A trans-activator of nuclear-encoded respiratory genes in animal cells. Genes & Dev. 4: 1023–1034.

Graham, F.L. and A.J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52: 456–467.

Hager, D.A. and R.R. Burgess. 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity. Results with the sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal. Biochem. 109: 76–86.

Hatefi, Y. 1985. The mitochondrial electron transport chain and oxidative phosphorylation system. Annu. Rev. Biochem. 54: 1015–1069.

Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26: 365–369.

Hochschild, A. 1992. Detecting cooperative protein-DNA interactions and DNA loop formation by footprinting. Methods Enzymol. 208: 343–361.

Kadonaga, J.T. and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. 83: 5889–5893.

Karim, F.D., L.D. Ureness, C.S. Thomas, M.J. Klemsz, S.R. Mckercher, A. Celada, C. Van Beveren, R.A. Mak, 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.

LaMarco, K.L. and S.L. McKnight. 1989. Purification of a set of cellular polyptides that bind to the purine-rich cis-regulatory element of herpes simplex virus immediate early genes. Genes & Dev. 3: 1372–1383.

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. Genes & Dev. 5: 975–990.

Nye, J.A., M. Petersen, C.V. Gunther, M.D. Jonsen, and B.J. Graves. 1992. Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. Genes & Dev. 6: 975–990.

Seto, E., Y. Shi, and T. Shenk. 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. Nature 354: 241–245.

Shapiro, D.J., P.A. Sharp, W.W. Wahl, and M.J. Keller. 1988. A high-efficiency HeLa cell nuclear transcription extract. DNA 7: 47–55.
Virbasius et al.

Smale, S.T. and D. Baltimore. 1989. The “initiator” as a transcription control element. Cell 57: 103–113.
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.
Virbasius, J.V. and R.C. Scarpulla. 1990. The rat cytochrome c oxidase subunit IV gene family: Tissue-specific and hormonal differences in subunit IV and cytochrome c mRNA expression. Nucleic Acids Res. 18: 6581–6586.
——. 1991. Transcriptional activation through ETS domain binding sites in the cytochrome c oxidase subunit IV gene. Mol. Cell. Biol. 11: 5631–5638.
Watanabe, H., T. Wada, and H. Handa. 1990. Transcription factor E4TF1 contains two subunits with different functions. EMBO J. 9: 841–847.
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.

J V Virbasius, C A Virbasius and R C Scarpulla

*Genes Dev.* 1993, 7:
Access the most recent version at doi:10.1101/gad.7.3.380

References
This article cites 26 articles, 13 of which can be accessed free at: http://genesdev.cshlp.org/content/7/3/380.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.