

NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators

Ching-man A. Virbasius, Joseph V. Virbasius, and Richard C. Scarpulla¹

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

Nuclear respiratory factor 1 (NRF-1) was first discovered as an activator of the cytochrome *c* gene and was subsequently found to play a broader role in nuclear-mitochondrial interactions. We have now cloned a HeLa cDNA encoding NRF-1 using degenerate oligomers derived from tryptic peptide sequences for PCR amplification. The cDNA-encoded protein was indistinguishable from the authentic HeLa cell factor on denaturing gels, displayed the expected NRF-1 DNA-binding specificity, and made the same guanine nucleotide contacts as HeLa NRF-1 on binding known NRF-1 recognition sites. Antiserum raised against the highly purified recombinant protein recognized the identical DNA-protein complex formed using either a crude nuclear fraction or nearly homogeneous HeLa NRF-1. Recombinant NRF-1 also activated transcription through specific sites from several NRF-1-responsive promoters, confirming both the transcriptional activity and specificity of the cDNA product. Portions of NRF-1 are closely related to sea urchin P3A2 and the erect wing (EWG) protein of *Drosophila*. Both are recently identified developmental regulatory factors. The region of highest sequence identity with P3A2 and EWG was in the amino-terminal half of the molecule, which was found by deletion mapping to contain the DNA-binding domain, whereas the carboxy-terminal half of NRF-1 was highly divergent from both proteins. The DNA-binding domain in these molecules is unrelated to motifs found commonly in DNA-binding proteins; thus, NRF-1, P3A2, and EWG represent the founding members of a new class of highly conserved sequence-specific regulatory factors.

[Key Words: Oxidative phosphorylation; nuclear respiratory factors; mitochondria; transcription]

Received July 27, 1993; revised version accepted October 1, 1993.

The vertebrate mitochondrion contains its own genome, along with the machinery required for its autonomous transcription, translation, and replication (Attardi and Schatz 1988; Clayton 1991; Wallace 1992). Mitochondrial DNA, however, has a coding capacity limited to only 13 respiratory chain polypeptides, and 2 ribosomal and 22 transfer RNAs. Most respiratory proteins and all of those necessary for maintenance and expression of the mitochondrial genome are encoded in the nuclear DNA. Thus, understanding the genetic control of mitochondrial function largely becomes a problem of identifying the nuclear genes involved and investigating potential mechanisms of regulated expression.

Investigations of nuclear-mitochondrial interactions in mammalian cells have led to the cloning of nuclear gene products required for mitochondrial DNA transcription and replication (Clayton 1991). The human and mouse genes for the RNA subunit of MRP endonuclease,

a ribonucleoprotein enzyme that is thought to cleave light-strand transcripts to form primers for heavy-strand DNA replication, have been cloned (Chang and Clayton 1989; Topper and Clayton 1990). A second nuclear gene product linking transcription and replication is mitochondrial transcription factor 1 (mtTF1), now called mtTFA (Xu and Clayton 1992). This factor recognizes the divergent heavy- and light-strand promoters to stimulate transcription initiation by mitochondrial RNA polymerase in vitro (Fisher et al. 1987). Because light-strand transcripts cleaved by mitochondrial RNA processing (MRP) endonuclease prime heavy-strand replication, mtTFA has the potential to modulate both transcription and replication of mitochondrial DNA (Clayton 1991). The importance of mtTFA to mitochondrial function in vivo is supported by the observation that a null mutation in the yeast counterpart to mtTFA (ABF2) results in the loss of mitochondrial DNA (Diffley and Stillman 1991). Interestingly, this phenotype can be rescued by expression of human mtTFA in yeast (Parisi et al. 1993).

Regulatory elements common to nuclear genes with

¹Corresponding author.

products that function in the mitochondria have also been described (Nagley 1991; Wallace 1993). Analysis of cytochrome *c* and cytochrome oxidase promoters has led to the identification of transcriptional activators designated as nuclear respiratory factors (NRF)-1 (Evans and Scarpulla 1989, 1990; Chau et al. 1992) and -2 (Virbasius and Scarpulla 1991; Virbasius et al. 1993). Functional NRF-1 sites have been found in genes encoding cytochrome *c* and at least one subunit each of respiratory complexes III, IV, and V (Evans and Scarpulla 1989, 1990; Chau et al. 1992), suggesting a role for the factor in the coordinate expression of respiratory chain subunits. NRF-1 may also participate in mitochondrial gene expression through its sequence-specific activation of genes encoding both the MRP RNA (Evans and Scarpulla 1990) and mtTFA (this paper; Virbasius and Scarpulla 1994). Activity of the proximal mtTFA promoter is highly dependent on NRF-1 in both transfected cells and in *in vitro* transcription assays (Virbasius and Scarpulla 1994). Similarly, expression of the gene encoding 5-aminolevulinate (5-ALA) synthase, the rate-limiting enzyme in the biosynthesis of heme for respiratory cytochromes, requires two NRF-1 recognition sites within its promoter region (Braidotti et al. 1993). These findings are consistent with an integrative role for NRF-1 in controlling nuclear-mitochondrial interactions in higher organisms. In addition, functional NRF-1-binding sites are present in the genes for tyrosine aminotransferase and the translation initiation factor eIF-2 α (Chau et al. 1992). Like 5-ALA synthase, these two proteins participate in the rate-limiting steps of their respective pathways of tyrosine catabolism and protein synthesis. Thus, NRF-1 may integrate a number of metabolic processes by regulating the genes encoding key enzymes.

As a prelude to molecular cloning, NRF-1-binding activity was purified over 30,000-fold to near homogeneity and was found to reside in a single polypeptide of 68 kD (Chau et al. 1992). Here, we use the sequences of tryptic peptides to obtain a cDNA clone that encodes a protein with the DNA-binding and transcriptional specificities expected for NRF-1. Deletion mapping of NRF-1 establishes that its DNA-binding domain coincides with a region of high sequence similarity with P3A2 (Calzone et al. 1991; Hoog et al. 1991) and erect wing (EWG) (Desimone and White 1993), two recently identified developmental regulatory factors. Thus, these proteins define a new family of sequence-specific regulators that share a conserved DNA-binding domain. On the basis of functional analysis of over a dozen NRF-1 sites, we predict that the NRF-1 DNA-binding domain participates in the expression of >50 mammalian genes of known sequence.

Results

Molecular cloning and overexpression of recombinant NRF-1

Tryptic peptides were derived from ~50 pmoles of purified HeLa NRF-1 (Chau et al. 1992). Two peptide peaks

from HPLC chromatography were chosen for microsequencing. NRF-1(72) yielded 26 residues (SMILEDLE-SALAEHAPAPQEVNSELPL) from a single homogenous peptide, whereas NRF-1(38) was a mixture of two peptides with a novel 8-residue sequence (VSWTQALR) derived from the secondary product. The major NRF-1(38) peptide was from Ku antigen, a known contaminant of DNA-binding proteins (Kadonaga 1991). A series of degenerate oligonucleotides was designed for PCR amplification of total cDNA prepared from HeLa poly(A)⁺ RNA. Two of these probes yielded a 269-nucleotide PCR product containing an open reading frame that was subsequently used as a probe to obtain a 3-kb HeLa cDNA clone. The putative NRF-1 cDNA had a 503-amino-acid open reading frame containing the complete sequence of the PCR product flanked by the two NRF-1 peptides (Fig. 1).

The cDNA product was overexpressed in *Escherichia coli* using an inducible T7 expression system (Studier et al. 1990). A transformant containing the NRF-1-coding region transcribed from an inducible T7 promoter gave a major 68-kD protein on induction (Fig. 2, lanes 2,3). The 68-kD mass of the induced protein is identical to that determined previously for highly purified HeLa NRF-1 (Chau et al. 1992) but is greater than the 54-kD mass predicted by the NRF-1 open reading frame. As only the open reading frame fragment was cloned into the expression vector, the mass of NRF-1 appears to be overestimated on denaturing gels. The induced 68-kD protein was purified to >95% purity from sonified extracts by ammonium sulfate precipitation (lane 4) and heparin-agarose fractionation (lane 5).

To determine whether the cDNA-encoded recombinant protein was the same as that present in DNA-protein complexes formed using HeLa cell NRF-1, goat antiserum was raised against the purified recombinant protein and tested for its ability to "supershift" NRF-1-DNA complexes in a gel-retardation assay. DNA-protein complexes of identical migration were formed with a labeled rat cytochrome *c* NRF-1 oligomer (RC4 - 172/ - 147) using crude nuclear extract (Fig. 3, lane 1), affinity-purified NRF-1 (lane 4), or recombinant NRF-1 (lane 7). In each case, the complex was supershifted with antiserum raised against recombinant NRF-1 (lanes 3,6,9) under conditions where preimmune serum had no effect (lanes 2,5,8). Variations in the antibody complexes formed with affinity-purified and recombinant proteins result from differences in the antigen-antibody ratios present in each reaction. The multiple complexes observed previously with binding of NRF-2 to its recognition site in the cytochrome oxidase subunit Vb (MCO5b + 13/ + 33) gene (Virbasius et al. 1993) were unaffected by the addition of anti-NRF-1 or preimmune serum to binding reactions (lanes 10-12). These results provide a direct link between the NRF-1 cDNA product and the NRF-1-binding activity present in HeLa cells.

DNA-binding specificity of recombinant NRF-1

The recombinant protein should display the same DNA-

ACCGCTGCCGCTTCTGAGT 19

20 TCCTGGTCCAGACTTTACACAGAAAAGTGTCTCAAAGGATATTTGTTTAATGAATGTGGTATGCTGACATTTAAACAGGACAAAACCTGTAGAACTTC 118

10

119 Met Glu Glu His Gly Val Thr Gln Thr Glu His Met Ala Thr Ile Glu Ala His Ala Val Ala Gln Gln Val Gln 193
 ATG GAG GAA CAC GGA GTG ACC CAA ACC GAA CAT ATG GCT ACC ATA GAA GCA CAT GCA CTG GCC CAG CAA GTG CAG 50

30

194 Gln Val His Val Ala Thr Tyr Thr Glu His Ser Met Leu Ser Ala Asp Glu Asp Ser Pro Ser Ser Pro Glu Asp 268
 CAG GTC CAT GTG GCT ACT TCT ACC GAG CAT AGT ATG CTG AGT GCT GAT GAA GAC TCC CCT TCT TCG CCC GAG GAC 50

70

269 Thr Ser Tyr Asp Asp Ser Asp Ile Leu Asn Ser Thr Ala Ala Asp Glu Val Thr Ala His Leu Ala Ala Ala Gly 343
 ACC TCT TAC GAT GAC TCA GAT ATA CTC AAC TCC ACA GCA GCT GAT GAG GTG ACA GCT CAT CTG GCA GCT GCA GGT 90

90

344 Pro Val Gly Met Ala Ala Ala Ala Val Ala Thr Gly Lys Lys Arg Lys Arg Pro His Val Phe Glu Ser Asn 418
 CCT GTG GGA ATG GCC GCT GCT GCT GCT GCA ACA GGA AAG AAA CCG AAA CCG CCT CAT GTA TTT GAG TCT AAT 110

110

419 Pro Ser Ile Arg Lys Arg Gln Gln Thr Arg Leu Leu Arg Lys Leu Arg Ala Thr Leu Asp Glu Tyr Thr Thr Arg 493
 CCA TCT ATC CCG AAG AGG CAA CAA ACA CCT TTG CTT CGG AAA CTT CGA GCC ACC TTA GAT GAA TAT ACT ACT CGT 130

130

494 Val Gly Gln Gln Ala Ile Val Leu Cys Ile Ser Pro Ser Lys Pro Asn Pro Val Phe Lys Val Phe Gly Ala Ala 568
 GTG GGA CAG CAA GCT ATT GTC CTC TCT ATC TCA CCC TCC AAA CCT AAC CCT GTC TTT AAA GTG TTT GGT GCA GCA 170

170

569 Pro Leu Glu Asn Val Val Arg Lys Tyr Lys Ser Met Ile Leu Glu Asp Leu Glu Ser Ala Leu Ala Glu His Ala 643
 CCT TTG GAG AAT GTG CTG GCT AAG AGC ATG ATC CTG GAA GAC CTG GAC TCT CCT CTG GCA GAA CAG GCC 190

190

644 Pro Ala Pro Gln Glu Val Asn Ser Glu Leu Pro Pro Leu Thr Ile Asp Gly Ile Pro Val Ser Val Asp Lys Met 718
 CCT GCG CCA CAG GAG GTT AAC TCA GAA CTG CCG CCT CTC ACC ATC GAC GGA ATT CCA GTC TCT GTG GAC AAA ATG 210

210

719 Thr Gln Ala Gln Leu Arg Ala Phe Ile Pro Glu Met Leu Lys Tyr Ser Thr Gly Arg Gly Lys Pro Gly Trp Gly 793
 ACC CAG GCC CAG CTT CCG GCA TTT ATC CCA GAG ATG CTC AAG TAC TCT ACA GGT CCG GGA AAA CCA GGC TGG GGG 230

230

794 Lys Glu Ser Cys Lys Pro Ile Trp Trp Pro Glu Asp Ile Pro Trp Ala Asn Val Arg Ser Asp Val Arg Thr Glu 868
 AAA GAA AGC TGC AAG CCC ATC TGG TGG CCT GAA GAT ATC CCC TGG GCA AAT GTC CCG AGT GAT GTC CCG ACA GAA 250

250

869 Glu Gln Lys Gln Arg Val Ser Trp Thr Gln Ala Leu Arg Thr Ile Val Lys Asn Cys Tyr Lys Gln His Gly Arg 943
 GAG CAA AAG CAG AGG GTT TCA TGG ACC CAG GCA CTA CCG ACC ATA GTT AAA AAC TGT TAT AAA CAG CAT GGG CCG 290

290

944 Glu Asp Leu Leu Tyr Ala Phe Glu Asp Gln Thr Gln Thr Gln Ala Thr Ala Thr His Ser Ile Ala His Leu 1018
 GAA GAC CTT TTG TAT GCC TTT GAA GAT CAG CAA AGC CAA ACA CAG GCC ACA GCC ACA CAT AGT ATA GCT CAT CTT 310

310

1019 Val Pro Ser Gln Thr Val Val Gln Thr Phe Ser Asn Pro Asp Gly Thr Val Ser Leu Ile Gln Val Gly Thr Gly 1093
 GTA CCA TCA CAG ACT GTA CTC CAG ACT TTT AGT AAC CCT GAT GGC ACT GTC TCA CTT ATC CAG GTT GGT ACG GGG 330

330

1094 Ala Thr Val Ala Thr Leu Ala Asp Ala Ser Glu Leu Pro Thr Thr Val Thr Val Ala Gln Val Asn Tyr Ser Ala 1168
 GCA ACA GTA GCC ACA TTG GCT GAT GCT TCA GAA TTG CCA ACC ACG GTC ACC GTT GCC CAA GTG AAT TAT TCT GCC 370

370

1169 Val Ala Asp Gly Glu Val Glu Gln Asn Trp Ala Thr Leu Gln Gly Gly Glu Met Thr Ile Gln Thr Thr Gln Ala 1243
 GTG GCT GAT GGA GAG GTC GAA CAA AAT TGG GCC ACG TTA CAG GGA GGT GAG ATC ACC ATC CAG ACG ACG CAA GCA 390

390

1244 Ser Glu Ala Thr Gln Ala Val Ala Ser Leu Ala Glu Ala Val Ala Val Ala Lys Ser Gln Glu Met Gln Gln Gly Ala 1318
 TCA GAG GCC ACC CAG GCC GTC GCA TCG TTG GCA GAG GCC CCA CTG GCA GCT TCT CAG GAG ATG CAG CAG GGA GCT 410

410

1319 Thr Val Thr Met Ala Leu Asn Ser Glu Ala Ala His Ala Val Ala Thr Leu Ala Glu Ala Thr Leu Gln Gly 1393
 ACA GTC ACT ATG GCC CTT AAC AGC GAA GCT GCC CCG CAT GCT GTC GCC ACC CTC GCT GAG GGC ACC TTA CAA GGT 430

430

1394 Gly Gly Gln Ile Val Leu Ser Gly Glu Thr Ala Ala Ala Val Gly Ala Leu Thr Gly Val Ser Asp Ala Asn Gly 1468
 GGG GGA CAG ATC GTC TTT TCT GGG GAA ACC GCA GCA Ala Ala Val Gly GCA CTT ACT GGA GTC CAA GAT GCT AAT GGC 470

470

1469 Leu Val Gln Ile Pro Val Ser Met Tyr Gln Thr Val Val Thr Ser Leu Ala Gln Gly Asn Gly Pro Val Gln Val 1543
 GTC GTC CAG ATC CCT GTC AGC ATG TAC CAG ACT GTG GTG ACC AGC CTC GCC CAG GCC AAC GGA CCA GTG CAG GTG 490

490

1544 Ala Met Ala Pro Val Thr Thr Arg Ile Ser Asp Ser Ala Val Thr Met Asp Gly Gln Ala Val Glu Val Val Thr 1618
 GCC ATG GCC CCT CTC ACC ACC AGG ATA TCA GAC AGC GCA CTC ACC ATG GAC GGC CAA GCT GTG GAG GTG GTG ACA 510

510

Leu Glu Gln End
 1619 TTG GAA CAG TCA

Figure 1. Nucleotide and predicted amino acid sequence of the HeLa NRF-1 cDNA 5'-untranslated and coding region. Amino acid sequences matching those of tryptic peptides from the purified NRF-1 protein are underlined. The sequence coinciding with the PCR product amplified from HeLa cDNA using primers derived from the NRF-1 peptide sequences is indicated with a bold underline. The complete 2970-nucleotide NRF-1 cDNA sequence has been submitted to GenBank under accession number L22454.

binding properties ascribed previously to NRF-1 (Evans and Scarpulla 1989, 1990; Chau et al. 1992). Recombinant NRF-1 was thus used for competition DNase I footprinting of the rat cytochrome *c* promoter region (Fig. 4). In the absence of competitor, the recombinant protein yielded several intense enhanced cleavages at the 5' and 3' ends of the footprint, with the absence of cleavages throughout the intervening protected region (Fig. 4, lane 3). This pattern is identical to that observed previously using preparations of HeLa NRF-1 (Evans and Scarpulla 1990). The footprint was eliminated by the inclusion of an excess of unlabeled oligonucleotides of previously characterized NRF-1 sites from nuclear genes with products that function in the mitochondria. These include cytochrome *c* (RC4, lane 4), cytochrome oxidase subunit VIc (COXVIc, lane 5), and mouse MRP RNA (mMRP, lane 6). Moreover, a sequence from the cytochrome *c*₁ gene with two mismatches from the NRF-1 consensus did not compete (hCC₁, lane 9), but a mutated derivative, active in both NRF-1 binding and transcriptional activity (Evans and Scarpulla 1990), did (hCC₁UP, lane 10). The unrelated cytochrome *c* ATF/CREB site (RC4 -281/-256, lane 11) served as a negative control.

In addition to the known NRF-1 sites, we observed strong similarities to the NRF-1 consensus in recently isolated genes encoding cytochrome oxidase subunit Vb (COXVb) [Basu and Avadhani 1991] and mtTFA (Tomimaga et al. 1992). Oligomers of each of these sites were

tested for their ability to stimulate the activity of a truncated cytochrome *c* promoter in transfected cells. The COXVb -109/-87 and mtTFA -73/-46 oligomers stimulated promoter activity 12.3±3.4-fold and 6.0±1.9-

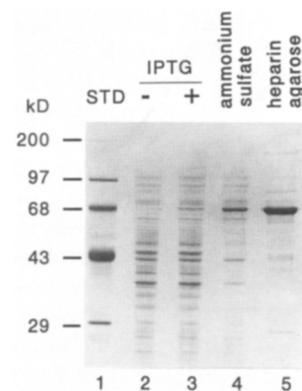


Figure 2. Expression and purification of recombinant NRF-1. Coomassie-blue stained SDS-PAGE of molecular mass standards (lane 1) and 50 µl of log-phase culture of *E. coli* strain BL21(DE3) transformed with the NRF-1-coding region in the pET3d expression vector uninduced (lane 2) or induced by the addition of 0.4 mM IPTG (lane 3). A lysate of an induced culture was purified by ammonium sulfate precipitation (lane 4) and fractionation of the pellet on a heparin-agarose column (lane 5). Lanes 4 and 5 contain equal amounts (4 µg) of total protein.

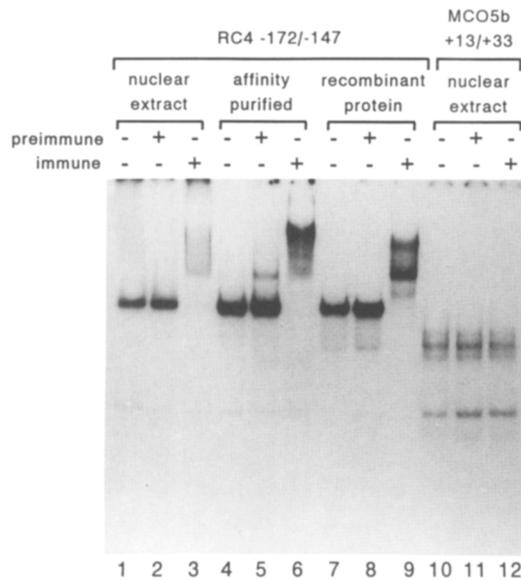


Figure 3. Recognition of HeLa NRF-1 by antiserum directed against the recombinant protein. Binding reactions contained 14 μ g of HeLa nuclear extract (lanes 1–3, 10–12), 12 ng of affinity-purified HeLa NRF-1 (lanes 4–6), or 20 ng of bacterial NRF-1 heparin–agarose fraction (lanes 7–9). Labeled oligonucleotides contained either an NRF-1-binding site from the rat cytochrome *c* gene (RC4 –172/–147, lanes 1–9) or an NRF-2-binding site from the mouse COXVb gene (Virbasius et al. 1993) (lanes 10–12). Following the binding reaction, 1 μ l of preimmune serum or goat antiserum was added to bacterially produced NRF-1, and the complexes were subsequently resolved on a native polyacrylamide gel.

fold, respectively, when cloned in *cis*, results similar to the value of 10.1 ± 3.1 -fold obtained with the cytochrome *c* NRF-1 site (RC4 –171/–147). Both also formed specific complexes with affinity-purified HeLa NRF-1 (not shown). The COXVb and mtTFA oligomers were also found to be specific competitors in the DNase I footprinting assay using recombinant NRF-1 (Fig. 4, lanes 7,8). These results further substantiate the binding specificity of the recombinant protein and indicate that COXVb and mtTFA genes are likely to have NRF-1-responsive promoters.

If the recombinant protein is NRF-1, it should contact DNA through specific guanine nucleotides spanning one turn of the DNA helix, as demonstrated using preparations of the HeLa protein (Evans and Scarpulla 1990; Chau et al. 1992). Recombinant NRF-1 was thus used for methylation interference footprinting of known sites from RC4, COXVc, and MRP RNA genes. The pattern of guanine nucleotide contacts in each case was indistinguishable from that obtained using HeLa NRF-1 and conforms to the consensus derived previously (Fig. 5). The RC4 site was known to deviate from the others by making additional downstream contacts (Evans and Scarpulla 1990). The same pattern is observed here with recombinant NRF-1. Therefore, the binding interactions between recombinant NRF-1 and cognate sites in several promot-

ers are identical to those observed with the HeLa protein, providing further support for the conclusion that the isolated cDNA encodes NRF-1.

Transcriptional activity and specificity of recombinant NRF-1

The binding of NRF-1 to its recognition site has been correlated with site-specific effects on promoter activity in transfected cells (Evans and Scarpulla 1989, 1990; Chau et al. 1992). The cloning of NRF-1 now affords an opportunity for a direct demonstration of its transcriptional activity and specificity. Promoter activation by NRF-1 was thus tested in an *in vitro* transcription assay using both a wild-type promoter from the rat cytochrome *c* gene (RC4CAT/–326) and the same template containing an insertional disruption of the NRF-1 site (RC4CAT/–326; LI–162/–159) that diminishes both NRF-1 binding and the activity of the transfected

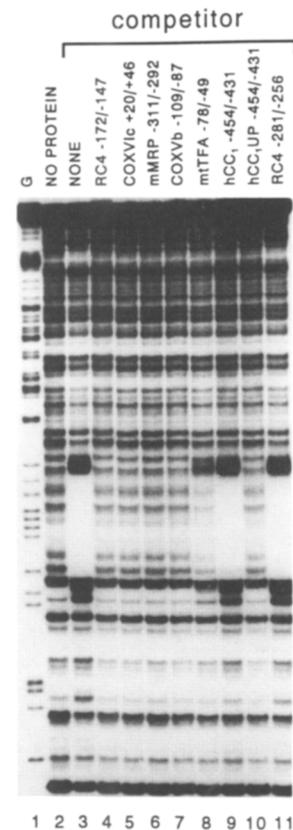


Figure 4. Binding of recombinant NRF-1 to the rat cytochrome *c* promoter region. An end-labeled RC4 promoter fragment containing the NRF-1-binding site was subjected to DNase I digestion following incubation in a mixture without added protein (lane 2) or with the addition of 20 ng of NRF-1 heparin–agarose fraction (lanes 2–11). Competitor oligonucleotides indicated above lanes 4–11 were added at a 200-fold molar excess before the addition of the labeled fragment. The extent of the NRF-1 footprint is indicated by the vertical bar at right. (G) G reaction of the labeled fragment.

RNA (4XmMRP) promoters were cloned into an RC4 vector deleted of sequences upstream from -66, and the resulting constructs were used for in vitro transcription. Compared with the vector with no insert (Fig. 6B, lanes 2-4), those with the RC4 (lanes 5-7) or mMRP (lanes 8-10) NRF-1 sites displayed a strong, dose-dependent increase in transcription in response to added recombinant NRF-1. In this case, no inhibition was observed at 0.4 μ g of NRF-1 because of the increased binding capacity of the promoter template for NRF-1. As with the intact cytochrome *c* promoter, initiation occurred at the same site used by the liver initiation complex in the synthesis of cytochrome *c* mRNA. These results establish that NRF-1 is a transcriptional activator that can function both in the proper promoter context and in a minimal promoter to direct the synthesis of high levels of accurately initiated transcripts.

NRF-1 has a new DNA-binding domain conserved in developmental regulatory factors

It was of interest to determine whether NRF-1 shares structural features with other proteins. A computer search revealed a region of extensive sequence similarity with two recently described developmental regulatory factors (Fig. 7). The first, P3A2, has been implicated in the correct expression of a cytoskeletal actin gene during sea urchin development (Calzone et al. 1991; Hoog et al. 1991). The second, the EWG gene product of *Drosophila*

melanogaster participates in both nervous system and flight muscle development (Desimone and White 1993). Binding to DNA has been demonstrated for P3A2 but not EWG, and neither has yet been shown to function directly in transcriptional activation.

Alignment of NRF-1 with P3A2 and EWG reveals a stretch of striking sequence conservation among all three proteins between NRF-1 residues 65 and 284 (Fig. 7). This region coincides with the novel DNA-binding domain identified previously for P3A2 (P3A2 residues 25-258) (Hoog et al. 1991) and corresponding here to NRF-1 residues 61-290. In contrast, the three proteins share little similarity in their carboxy-terminal halves or in an amino-terminal extension present in NRF-1 and EWG.

To determine whether the highly conserved region coincided with the NRF-1 DNA-binding domain, a deletion series of truncated NRF-1 molecules (summarized diagrammatically in Fig. 8A) was expressed by in vitro transcription and translation, and the products were assayed for binding to radiolabeled RC4 -172/-147. As shown in Figure 8B, lane A, the intact cDNA yielded a translation product migrating at 68 kD. This protein was unaltered by deletion of the 3'-untranslated region to a position just downstream from the predicted NRF-1 translational terminator (lane B), confirming that the translation product is derived from the NRF-1 open reading frame. To demonstrate that the 68-kD translation product had the correct binding specificity, it was tested

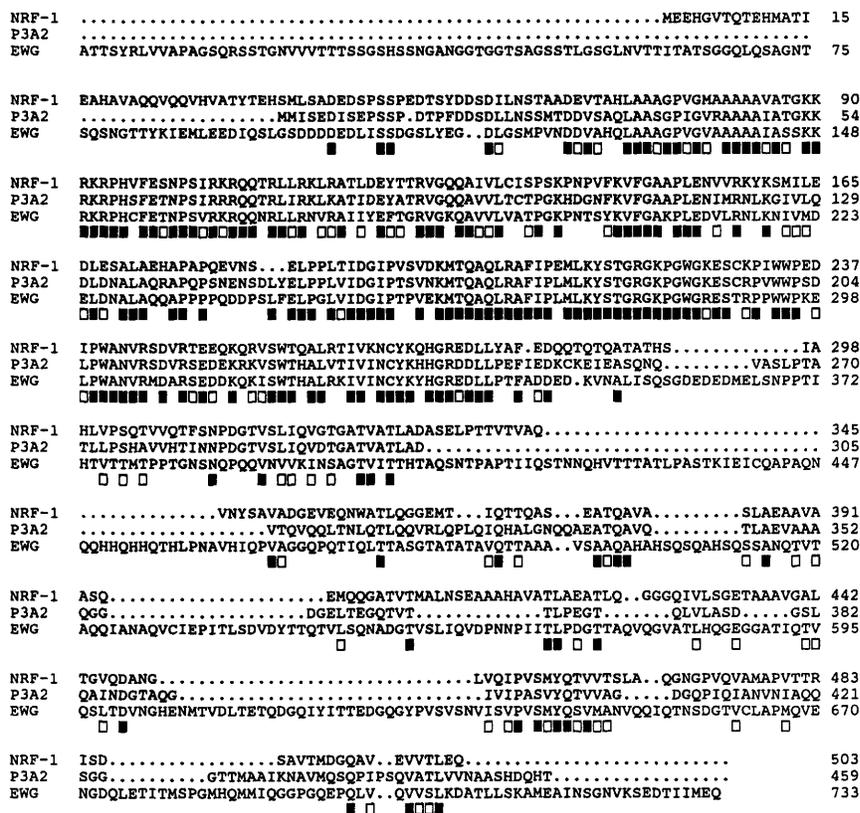


Figure 7. Alignment of NRF-1 protein sequence with those of developmental regulatory factors P3A2 and EWG. Sequences of human NRF-1, sea urchin P3A2, and *Drosophila* EWG were aligned using the GAP program of the Genetics Computer Group (program manual, v.7, 1991). (■) Residues identical in all three proteins; (□) positions where all three proteins contain similar (conservative) amino acid substitutions grouped as follows: (A S T); (D E); (N Q); (R K); (I L M V); (F Y W). Dots denote gaps introduced for optimum alignment.

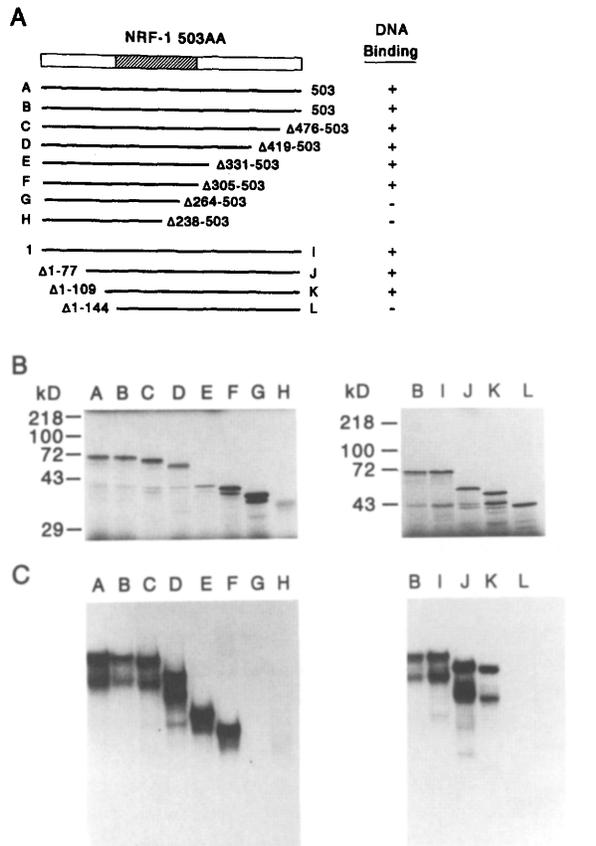


Figure 8. Deletion mapping of the NRF-1 DNA-binding domain. (A) Schematic representation of in vitro-translated proteins tested for DNA binding. The shaded box represents the outer limits of the region required for DNA-binding activity. Construct A includes the complete cDNA sequence. In construct B all but 35 nucleotides of the 3'-untranslated region is deleted. C-H represent carboxy-terminal deletions of the residues indicated and were generated by either cleavage at a restriction endonuclease site in the native sequence (C,D,E,H) or insertion of a synthetic translation terminator (F,G). In I the entire 5'-untranslated region, which includes several potential upstream initiators, was deleted, and J,K and L represent amino-terminal deletions of the indicated residues and restoration of the initiator ATG by the addition of an *Nco*I linker. Activity of the proteins in DNA-binding assays is summarized by + or - (right). (B) SDS-PAGE showing [³⁵S]methionine-labeled translation products corresponding to the constructions diagrammed in A. Positions of molecular mass standards are at the left of each panel. (C) Electrophoretic mobility shift assay for binding activities of deletion mutants. Translation extracts containing the proteins diagrammed in A were incubated with end-labeled RC4 -172/-147 oligonucleotide, and the products were resolved on native acrylamide gels.

for binding to known NRF-1-binding sites from nine different genes (Fig. 9). For each site, the major DNA-protein complex migrated with the complex formed with affinity-purified NRF-1 (lanes 3-5). Minor discrepancies result from differences in lengths of the labeled oligomers. All NRF-1 complexes were competitively displaced by an excess of unlabeled RC4 -172/-147 but

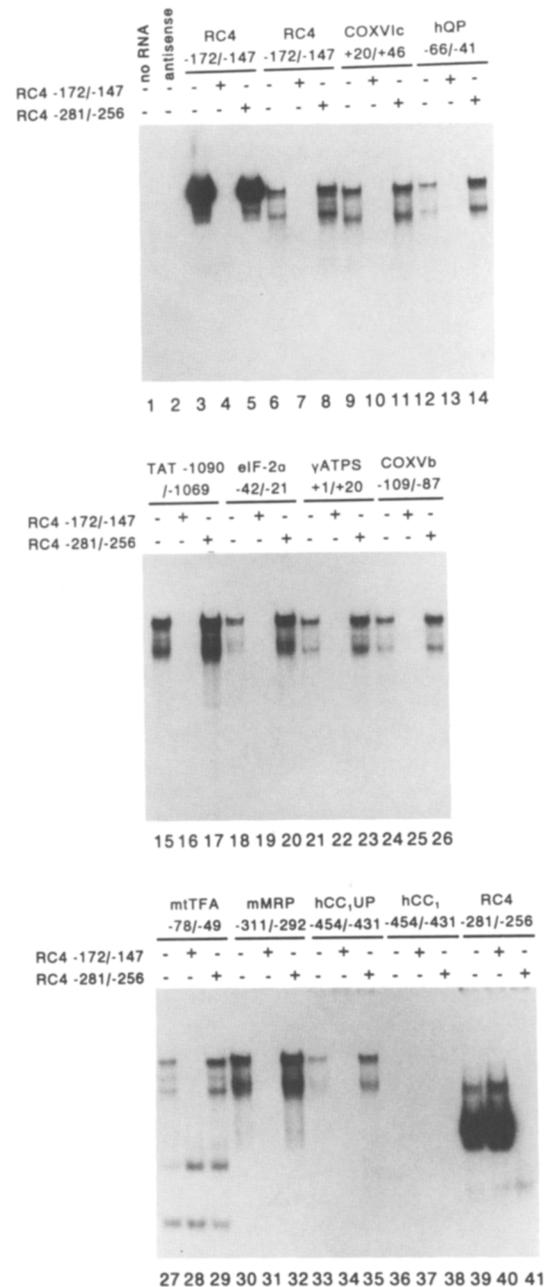


Figure 9. Specific binding of in vitro-transcribed and -translated NRF-1 to known NRF-1 recognition sites. Binding reactions contained 5 ng of affinity-purified HeLa NRF-1 (lanes 3-5) or 2 μ l of wheat germ lysate without added RNA (lane 1), with RNA transcribed in vitro from an antisense NRF-1 template in pSG5 (lane 2), or with RNA transcribed in the sense orientation (lanes 6-41). End-labeled oligonucleotides (10 fmole/lane) contained the NRF-1-binding sites designated above each panel. Lanes 1 and 2 also contained the labeled RC4 -172/-147 oligonucleotide. Binding reactions were carried out in the presence (+) or absence (-) of a 200-fold excess of unlabeled specific competitor (RC4 -172/-147) or a negative control oligonucleotide containing the rat cytochrome c ATF/CREB site (RC4 -281/-256). DNA-protein complexes were resolved on 5% acrylamide native gels.

not by an excess of RC4 -281/-256 ATF/CREB oligomer. The faster migrating complex appears to result from a truncated product formed during *in vitro* synthesis, as indicated by the elimination of the lower complex by carboxy-terminal deletion to residue 305 (Fig. 8C, lane F).

Having confirmed the correct binding specificity, the carboxy-terminal deletions (Fig. 8A, constructs A-H) were expressed (Fig. 8B, lanes A-H) and assayed for DNA-binding activity (Fig. 8C, lanes A-H). DNA binding was unaffected until residues between 305 and 264 were removed (lanes F,G). Likewise, when the amino-terminal deletions (Fig. 8A, constructs I-L) were expressed (Fig. 8B, lanes I-L) and assayed (Fig. 8C, lanes I-L), binding was lost on removal of residues 109-144 (lanes K,L). The precise deletion of the 5'-untranslated region in construct I removes several potential initiation codons without affecting the translated product, further confirming the identity of the NRF-1 reading frame. The carboxy-terminal boundary of the DNA-binding domain, determined here between NRF-1 residues 264 and 305, compares favorably with that determined previously for P3A2 between NRF-1 residues 255 and 290. The amino-terminal boundary determined here between residues 109 and 144 is somewhat downstream from the P3A2 boundary between NRF-1 residues 61 and 126 but is overlapping in the region between residues 109 and 126. These results establish that the major region of sequence similarity among these proteins resides in their DNA-binding domains. Thus, NRF-1, P3A2, and EWG define a new family of regulatory factors that share a highly conserved DNA-binding motif.

Discussion

Identification of the cDNA-encoded product as NRF-1

Purification and molecular cloning of NRF-1 were undertaken as a requisite for further investigating its structural characteristics and biological functions. Previously, we had purified NRF-1 >30,000-fold from HeLa nuclear extracts and demonstrated that a single 68-kD polypeptide accounted for specific binding to the known NRF-1 sites (Chau et al. 1992). The tryptic peptide sequences described here were derived from ~50 pmoles of the protein purified from >200 liters of HeLa cells.

The evidence presented here is consistent with the isolated cDNA encoding NRF-1. Both peptides derived from the purified protein were encoded in the cDNA, the expressed product of which migrated identically to HeLa NRF-1 on denaturing gels. Interestingly, the masses of NRF-1, P3A2, and EWG were all overestimated on denaturing gels by 30-50%, possibly reflecting shared structural features. Recombinant NRF-1 also binds specifically to the known NRF-1 sites through characteristic guanine contacts encompassing a single helical turn. Although the binding site is palindromic, the protein appears to bind as a monomer. Heterodimeric DNA-protein complexes were not detected when intact NRF-1 was mixed with derivatives truncated at either carboxyl

or amino termini, and the mass of HeLa NRF-1 was estimated at ~52 kD by glycerol gradient centrifugation in both the presence and absence of its binding site (C. Virbasius, unpubl.). Finally, *in vitro* transcription experiments unequivocally establish that the recombinant protein has the transcriptional activity and specificity expected for NRF-1. Taken together, the results presented here allow us to conclude with reasonable certainty that the cloned cDNA encodes NRF-1. However, given the existence of families of related transcription factors, it still remains a formal possibility that the true biological activity results from a protein that has escaped our detection.

Conservation of the NRF-1 DNA-binding domain in P3A2 and EWG

The striking conservation of the NRF-1 DNA-binding domain in P3A2 and EWG suggests that these proteins constitute a new family of regulatory factors with diverse functions in eukaryotic development. The P3A2 DNA-binding domain coincides with that defined here for NRF-1 and with the region of highest sequence conservation among all three factors. Although NRF-1 and P3A2 are acidic proteins with predicted isoelectric points of 4.71 and 5.49, respectively, nearly all of the lysine and arginine residues (33/34 for NRF-1 and 34/37 for P3A2) are clustered into two sequence blocks within the most highly conserved regions of the DNA-binding domain. The sequence between NRF-1 residues 89 and 160 (Fig. 7) is 25% lysine plus arginine and has 85% sequence conservation (identical plus similar residues) with P3A2. Likewise, the NRF-1 sequence between residues 199 and 274 is 20% lysine plus arginine and has 91% sequence conservation with P3A2.

In keeping with this structural conservation, the P3A2 recognition sites strongly resemble those for NRF-1. Both proteins make major groove contacts through alternating GC base pairs, and high-affinity binding occurs through a tandem repeat of the T/CGCGCA motif (Evans and Scarpulla 1990; Calzone et al. 1991). An apparent difference is that P3A2 can bind a monomer of this sequence at reduced affinity, whereas stable binding of NRF-1 requires a tandem direct repeat of this half-site (Table 1). No NRF-1 binding was detected to sequences from the cytochrome *c*₁ (hCC₁ -454/-431) and COX-VIc (COXVIc -46/-20) genes containing perfect NRF-1 half-sites (Evans and Scarpulla 1990). Two nucleotide changes in the hCC₁ site that restore the direct repeat (hCC₁UP -454/-431) also restore binding by NRF-1. This is confirmed here for these hCC₁ sites using the recombinant protein (Figs. 4 and 9), making it unlikely that NRF-1 would bind with high affinity to several of the P3A2 target sites. Also, in the highest affinity P3A2-binding sites, the half-site motifs are separated by intervening nucleotides and, in one case, are rotated by one-half helical turn (Calzone et al. 1991). These features have not been observed in the known NRF-1 recognition sites. It should be noted that P3A2 has been proposed to be a negative regulator of transcription through its dis-

Table 1. *NRF-1-binding sites*

| Gene name ^a | Sequence | Location ^b |
|---|--------------------|-----------------------|
| I. Functional binding sites | | |
| rat cytochrome <i>c</i> ¹ | CCCGCATGCGCG | -166/-155 |
| human cytochrome <i>c</i> ² | CCAGCATGCGCG | -169/-158 |
| rat cytochrome <i>c</i> oxidase subunit VIc ² | CACGCATGCGCA | +39/+28 |
| mouse cytochrome <i>c</i> oxidase subunit Vb ³ | CGCACATGCGCA | -92/-103 |
| bovine cytochrome <i>c</i> oxidase subunit VIIa ⁴ | TGCGCAAGCGCC | -11/-22 |
| human ubiquinone binding protein ² | TGCGCAGGCGCA | -53/-64 |
| bovine ATP synthase γ -subunit ⁵ | CGCGCACGCGCG | +4/+15 |
| mouse MRP RNA ² | TGCGCACGCGCA | -308/-297 |
| human MRP RNA ² | CGCGCACGCGCA | -293/-282 |
| human mitochondrial transcription factor A | GGCGCAGGCGCG | -59/-70 |
| rat 5-aminolevulinic synthase ⁶ | CGCGCACGCGCA | -77/-88 |
| | AGCGCATGCGCA | -59/-48 |
| rat tyrosine aminotransferase ⁵ | TGCACATGCGCA | -1085/-1074 |
| human eukaryotic initiation factor 2 α -subunit ⁵ | TCCGCATGCGCG | -37/-26 |
| Consensus | P yGCGCA P yGCGCPu | |
| II. Potential binding sites identified in GenBank/EMBL data base^c | | |
| A. Metabolic enzymes | | |
| human arylsulfatase A ⁷ | CGAGCACGCGCA | -71/-61 |
| human branched chain α -keto acid dehydrogenase ⁸ | TCCGCATGCGCA | -149/-138 |
| human carbonyl reductase ⁹ | CGCGCAGGCGCA | -46/-57 |
| human protein disulfide isomerase/prolyl 4-hydroxylase β ¹⁰ | CGCGCACGCGCC | -79/-90 |
| human α -enolase ¹¹ | CGCGCAGGCGCA | -734/-723 |
| rat glutamate dehydrogenase ¹² | CGAGCATGCGCA | -227/-216 |
| human steroid 5- α -reductase ¹³ | TGCGCACGCGCG | -269/-280 |
| mouse ornithine decarboxylase ¹⁴ | TGCGCACGCGCA | -266/-277* |
| human ornithine decarboxylase ¹⁵ | CGCGCACGCGCA | -333/-344* |
| rat ornithine decarboxylase ¹⁶ | CGCGCACGCGCA | -268/-279* |
| rat fatty acid synthase ¹⁷ | CGCGCAAGCGCG | -610/-621 |
| rat Na ⁺ /K ⁺ ATPase α -1 subunit ¹⁸ | CGCGCACGCGCG | -562/-549* |
| human calcium-activated neutral protease ¹⁹ | TGCGCATGCGCA | -18/-7 |
| human cathepsin D ²⁰ | GGCGCACGCGCA | -175/-186 |
| B. Signal transduction | | |
| mouse GM-CSF ²¹ | CACGCACGCGCG | -981/-992 |
| mouse hepatocyte growth factor-like protein ²² | CGCGCACGCGCA | -1014/-1003* |
| rat dopamine D1 receptor ²³ | CACGCACGCGCA | -1066/-1077 |
| human insulin receptor ²⁴ | GGCGCACGCGCG | -1036/-1025 |
| human insulin-like growth factor receptor ²⁵ | CGCGCACGCGCC | -29/-40 |
| human interferon receptor ²⁶ | CGCGCACGCGCC | -85/-96 |
| | TGCGCACGCGCT | +1/-11 |
| human cyclophilin ²⁷ | CGCGCACGCGCG | -173/-162* |
| human lipoprotein receptor-like protein ²⁸ | GGCGCATGCGCA | -151/-140 |
| rat calmodulin III ²⁹ | CGCGCACGCGCG | -209/-220* |
| | CGCGCACGCGCA | -197/-208* |
| human calretinin ³⁰ | GGCGCAGGCGCA | -53/-42 |
| human Go- α ³¹ | AGCGCACGCGCG | -835/-846 |
| human ADP ribosylation factor 1 ³² | GGCGCAGGCGCA | -215/-204 |
| human protein phosphatase 2A α ³³ | CGCGCAGGCGCA | -532/-543 |
| mouse cyclic nucleotide phosphodiesterase ³⁴ | TGCGCACGCGCC | -79/-90 |
| C. Chromosome maintenance and nucleic acid metabolism | | |
| human DNA polymerase α ³⁵ | TGCGCAAGCGCA | -235/-246 |
| human topoisomerase I ³⁶ | CGCGCAGGCGCA | -19/-8 |
| human H1 RNA ³⁷ | GGCGCACGCGCG | -145/-156 |
| human hnRNP core protein A1 ³⁸ | TGCGCAGGCGCA | -558/-569* |
| | TACGCATGCGCA | -521/-510 |
| | TGCGCAGGCGCA | -398/-401* |
| mouse S16 ribosomal protein ³⁹ | TGCGCAGGCGCA | -35/-46* |
| mouse histone H2a.2 ⁴⁰ | CGAGCACGCGCG | -210/199 |
| mouse histone H3 ⁴¹ | CGCACGCGCA | -40/-51 |

(Table 1 continued on following page.)

Table 1. NRF-1-binding sites (Continued)

| Gene name ^a | Sequence | Location ^b |
|--|--------------|-----------------------|
| human RCC1 ⁴² | TGCGCACGCGCA | -78/-67* |
| human cdc2 ⁴³ | CCAGCATGCGCA | -717/-706 |
| D. Other | | |
| human bcl-2 ⁴⁴ | CACGCACGCGCA | -889/-900 |
| human GADD153 growth/DNA damage inducible gene ⁴⁵ | CGCGCACGCGCA | -175/-186* |
| hamster GADD153 ⁴⁶ | CGCGCACGCGCA | -365/-354* |
| human 86-kD heat shock protein ⁴⁷ | CGCGCAGGCGCA | -286/-257 |
| human synapsin I ⁴⁸ | TGCGCACGCGCC | -241/-252 |
| rat synapsin I ⁴⁸ | TGCGCACGCGCC | -240/-251 |
| mouse myb proto-oncogene ⁴⁹ | GGCGCACGCGCC | -295/-306 |

^aThe search was confined to primate and rodent genes. For part I, references are given for the demonstration of NRF-1 binding. For part II, publications of the gene sequence are cited.

^bCoordinates are given relative to the first transcription start site, if known. Otherwise the authors' numbering system is followed. In some cases, the reverse complement of the published sequence is given to conform to the consensus. Sites identical to those of known function are indicated by an asterisk (*).

^cRodent and primate sequences in the GenBank (release 76) and EMBL data bases (release 34) were searched with the Findpatterns program (Genetics Computer Group Manual, 1991), allowing one mismatch to the consensus given. Only mismatches found in the known sites in part I were allowed in a further screen of the identified sequences. Furthermore, only sites in upstream regions of published genomic sequences were included in the table.

References: ¹(Evans and Scarpulla 1989), ²(Evans and Scarpulla 1990), ³(this work), ⁴(unpubl.), ⁵(Chau et al. 1992), ⁶(Braidotti et al. 1993), ⁷(Kreysing et al. 1990), ⁸(Mitsubuchi et al. 1991), ⁹(Forrest et al. 1991), ¹⁰(Tasanen et al. 1992), ¹¹(Giallongo et al. 1990), ¹²(Das et al. 1993), ¹³(Jenkins et al. 1991), ¹⁴(Katz and Kahana 1988), ¹⁵(Moshier et al. 1992), ¹⁶(Wen et al. 1989), ¹⁷(Amy et al. 1990), ¹⁸(Yagawa et al. 1990), ¹⁹(Miyake et al. 1986), ²⁰(Cavaillès et al. 1993), ²¹(Miyatake et al. 1985), ²²(Degen et al. 1991), ²³(Zhou et al. 1992), ²⁴(Tewari et al. 1989), ²⁵(Mamula and Goldfine 1992), ²⁶(Lutfalla et al. 1992), ²⁷(Haendler and Hofer 1990), ²⁸(Kutt et al. 1989), ²⁹(Nojima 1989), ³⁰(Parmentier and Lefort 1991), ³¹(Tsukamoto et al. 1991), ³²(Lee et al. 1992), ³³(KheW-Goodall et al. 1991), ³⁴(Kurihara et al. 1990), ³⁵(Pearson et al. 1991), ³⁶(Kunze et al. 1991), ³⁷(Baer et al. 1990), ³⁸(Biamonti et al. 1989), ³⁹(Wagner and Perry 1985), ⁴⁰(Hurt et al. 1989), ⁴¹(Sittman et al. 1983), ⁴²(Furuno et al. 1991), ⁴³(Ku et al. 1993), ⁴⁴(Adachi and Tsujimoto 1990), ⁴⁵(Park et al. 1992), ⁴⁶(Luethy et al. 1990), ⁴⁷(Walter et al. 1989), ⁴⁸(Sauerwald et al. 1990), ⁴⁹(Bender and Kuehl 1986).

placement of a zinc finger protein that binds the same sequence (Hoog et al. 1991). In contrast, NRF-1 clearly functions as a positive activator of transcription. Thus, it remains to be determined whether the structural conservation between P3A2 and NRF-1 in their DNA-binding domains will be precisely reflected in their binding and transcriptional specificities.

The EWG protein is required for viability of *Drosophila* embryos and for the proper development of the embryonic nervous system (Desimone and White 1993). Its molecular mass of 116 kD on denaturing gels is greater than that observed for P3A2 (62 kD) and NRF-1 (68 kD) and exceeds the mass predicted by its amino acid sequence (77 kD). Like P3A2, the sequence conservation with EWG is largely confined to the NRF-1 DNA-binding domain. Although its nuclear localization and structural conservation with P3A2 and NRF-1 are consistent with a function in gene regulation, there is as yet no evidence for DNA-binding or transcriptional effects, nor have potential target genes been identified.

NRF-1 and the nuclear control of mitochondrial function

The mitochondrion serves to compartmentalize diverse cellular metabolic systems largely regulated by enzymes encoded in the nuclear DNA. The apparatus for electron

transport and oxidative phosphorylation is unique in that both nuclear and mitochondrial genomes contribute protein subunits (Attardi and Schatz 1988; Clayton 1991; Wallace 1992). The sole purpose of the mitochondrial genetic system is to complement the contribution of nuclear genes in maintaining respiratory function. Such interplay between the two genomes might involve novel pathways of intracellular communication.

One possibility is that NRF-1 may help to coordinate the expression of respiratory chain subunits with components of the mitochondrial transcription and replication machinery. Such a model is consistent with the finding of functional NRF-1 recognition sites in genes encoding respiratory subunits, the MRP RNA (Evans and Scarpulla 1990; Chau et al. 1992) and mtTFA. The latter two have the capability of communicating changes in nuclear gene expression to the mitochondria through their effects on mitochondrial DNA replication and transcription. In keeping with this hypothesis, we have recently established that the proximal promoter for the mtTFA gene is almost completely dependent on a NRF-1 recognition site for its activity both in transfected cells and in an in vitro transcription assay using recombinant NRF-1 (Virbasius and Scarpulla 1994). NRF-1 control over mitochondrial function is also supported by the recent observation that the activity of 5-ALA synthase gene promoter is highly dependent on tandem NRF-1 recognition sites (Braidotti et al. 1993). Thus, NRF-1 con-

trol of this nuclear gene may serve to regulate the levels of the heme cofactor required by the respiratory cytochromes encoded by both genomes. These observations constitute a compelling case for an important integrative function for NRF-1 in communication between nuclear and mitochondrial genetic compartments.

Other regulatory functions for NRF-1 in coordinate gene expression

Our interest in nucleus-encoded mitochondrial functions, along with the remarkable consistency in the appearance of NRF-1 sites in the majority of genes in this category, has led to a direct functional characterization of these sites (Table 1). However, the identification of NRF-1 sites in the tyrosine aminotransferase and eIF-2 α genes suggested a broader integrative function for NRF-1 (Chau et al. 1992) and prompted a systematic search for potential binding sites in published gene sequences (Table 1). The sites listed have only a single mismatch with the consensus derived from the 14 tested binding sites; and in each case, the mismatch is known to be allowed at that position in the functional sites. These imposed constraints in selecting potential binding sites make it likely that the genes containing these sites are targets for NRF-1. In fact, 14 of the 48 putative target genes in Table 1 have NRF-1 sites identical to those of known function. It should be emphasized, however, that the effects of NRF-1 on basal promoter activity are influenced by promoter context. For example, mutation of the NRF-1 sites in the cytochrome *c* and COXVIc genes (Evans and Scarpulla 1989, 1990) results in a more modest effect on promoter activity than mutation of the NRF-1 sites in the mtTFA (Virbasius and Scarpulla 1994) or 5-ALA synthase genes (Braidotti et al. 1993). Also, the conservation of the NRF-1 DNA-binding domain in P3A2 and EWG suggests that this domain may be conserved in a mammalian family of related factors that mediate different biological functions through similar recognition sites. Thus, a rigorous analysis ultimately requires an evaluation of the NRF-1 sites within the proper promoter context and the identification of the cognate activator protein.

With these caveats in mind, some interesting observations emerge from Table 1. In addition to the cytochrome *c* and MRP RNA genes, there are several genes (ornithine decarboxylase, GADD153, and synapsin I) where the NRF-1 site is conserved in a similar location in different species. A majority of the genes are ubiquitously expressed, consistent with the wide distribution of the NRF-1-binding activity. In cases where ubiquitous and tissue-specific members of a gene family exist (cytochrome *c*, 5-ALA synthase, enolase, and the Na/K ATPase), NRF-1 sites are detected only in the widely expressed gene, suggesting involvement of NRF-1 in general, rather than tissue-specific cellular functions. Among a variety of metabolic enzymes encoded by these genes, several (ornithine decarboxylase and the branched-chain α -keto acid dehydrogenase) catalyze the rate-limiting step of their pathways (polyamine synthe-

sis and branched-chain amino acid catabolism, respectively). This is also true of tyrosine aminotransferase, 5-ALA synthase, and eIF-2 α , supporting a role for NRF-1 in integrating a variety of metabolic pathways by modulating the expression of a key activity. Putative NRF-1 sites in genes encoding a number of receptors and components of signal transduction networks may indicate a role in the establishment or maintenance of regulatory systems that influence these or other cellular functions. Also prominent in the list are genes involved in chromosome maintenance and nucleic acid metabolism. This may reflect a requirement for coordinating the expression of the replication, transcription, and translation machinery with organelle biogenesis under certain conditions. Similarly, NRF-1 sites are found in genes that may be directly involved in cell cycle regulation (*cdc2*, *RCC1*) or are regulated by cell growth (ornithine decarboxylase, DNA polymerase- α , and *GADD153*). Maintenance of mitochondria might be expected to require sensitivity to proliferative signals, and it is tempting to speculate that NRF-1 may function in transducing such signals. Thus, although the best defined biological role for NRF-1 is in the nuclear control of mitochondrial function, the NRF-1 protein or related proteins having the NRF-1 DNA-binding domain may have the potential for integrating diverse functions required for cell maintenance, growth, and proliferation.

Materials and methods

Purification and amino acid sequencing of NRF-1

DEAE-agarose and heparin-agarose fractionation of HeLa nuclear extracts have been described (Virbasius et al. 1993), except NRF-1 fractions were eluted with HEPES-D, 0.45 M KCl, diluted to 0.1 M KCl with HEPES-D, 0.1% NP-40, and loaded onto a NRF-1-specific affinity column as described (Chau et al. 1992). Affinity-purified NRF-1 was isolated by SDS-PAGE, transferred to nitrocellulose, and the NRF-1 band was identified by Ponceau-S staining (Aebersold et al. 1987). In situ tryptic digestion and peptide sequencing was performed by William S. Lane (Harvard Microchemistry Facility, Cambridge MA).

Amplification of NRF-1 sequence and cDNA library screening

One sense primer and two sets of antisense primers were used to amplify the NRF-1-coding sequence. One set of primers, 5'-GCIGATGCICCCIGCICCAIGAIGTAACTC-3' derived from the NRF-1(72) peptide and 5'-GCYTGNGTCCANGANAC-3' derived from the NRF-1(38) peptide, yielded a PCR product. Briefly, cDNA was synthesized using AMV reverse transcriptase (Promega) with 2 μ g of oligo(dT)-primed HeLa poly(A) RNA in a total volume of 20 μ l. The product (2 μ l) was mixed with two different pairwise combinations of sense and antisense primers and amplified with AmpliTaq DNA polymerase (Perkin-Elmer Cetus) for 50 cycles (94°C for 1 min, 50°C for 2 min, 72°C for 2 min). The products were ligated to M13mp18 for sequencing. A 269-bp PCR product, encoding portions of the two NRF-1 peptides, was subcloned into pGEM4 Blue. The insert was labeled by nick translation for screening a HeLa cDNA library in λ ZAPII [a gift of Dr. R. Morimoto, Northwestern University, Evanston, IL]. The 3-kb insert of one of two positive

phages was subcloned into pGEM7zf(+) (Promega) using the flanking *XhoI* and *XbaI* sites in the phage to generate pGEM7zf-NRF1. Subclones in M13 were sequenced on both strands using Sequenase (U.S. Biochemical).

Plasmid constructions

The full-length NRF-1 cDNA was put into the pSG5 expression vector (Green et al. 1988) in the sense orientation by ligating the upstream *EcoRI*-*PstI* fragment and downstream *PstI*-*BamHI* fragment of pGEM7zf-NRF1 into the *EcoRI* and *BamHI* sites of pSG5 to produce pSG5NRF1/1-2970. The antisense construction was generated by ligating the *XhoI*-*BamHI* fragment into the same sites of a pSG5 vector modified by insertion of a *XhoI* linker at the *EcoRI* site. pSG5NRF1/1-1662 (construct B in Fig. 8A) was generated by addition of a *BamHI* linker at the *AccI* site 35 bp downstream of the termination codon, removing the 3'-untranslated region. pSG5NRF1/1-1030 and pSG5NRF1/1-908 (constructs F and G in Fig. 8A) were generated by exonuclease III digestion from the *AccI* site, followed by addition of an *Asp718I* linker and cloning into a pGEM7zf(+) containing a synthetic universal translation terminator (Pharmacia) in its *SmaI* site. Deleted fragments were recloned to pSG5 using flanking *PstI* and *BamHI* sites. Amino-terminal truncations were generated either by PCR cloning or restriction enzyme cleavages. Briefly, pSG5NRF1/348-1662 (construct J in Fig. 8A) was generated by PCR using a sense primer, CCCATGGGAATGGCCGC, and an antisense primer, CCACGGCAGAATAATTC, matching sequence downstream of the natural *EcoRV* site. The PCR product was cut with *NcoI* and *EcoRV*, and the 480-bp fragment was cloned into pGEM5Zf(+) (Promega). An *EcoRI* linker was added at an adjacent *EagI* site. *EcoRI* and *HincII* digestion of this vector released a fragment that was then ligated with the *HincII*/*BamHI* fragment of pSG5NRF1/1-1662 into the *EcoRI* and *BamHI* sites of pSG5. pSG5NRF1/444-1662 and pSG5NRF1/551-1662 (constructs K and L in Fig. 8A) were generated by digestion of pSG5NRF1/1-1662 with *AflIII* and *DraI*, respectively, followed by the addition of *NcoI* linkers. Recloning into pSG5 was the same as described for pSG5NRF1/349-1662. pSG5NRF1/119-1662 (construct I in Fig. 8A) was generated by ligating the 221-bp *NcoI*-*PstI* fragment from pET3dNRF1 (see below) into pGEM5zf and cloning back into pSG5.

In vitro transcription, translation, and mobility shift assay

The pSG5 vectors described above include a T7 promoter upstream of the cloning site. To generate runoff transcription templates, pSG5NRF1/1-2970 was linearized with *BamHI* to generate the full-length (construct A in Fig. 8A), *NcoI* (C, Δ 476-503), *EcoNI* (D, Δ 419-503), *BglII* (E, Δ 331-503), or *EcoRV* (H, Δ 238-503). The other carboxy-terminal and all amino-terminal deletions were linearized with *BamHI*. In vitro transcription was performed by using T7 polymerase (Promega) and RNA translated in wheat-germ extract (Promega). Reactions contained unlabeled methionine for use in mobility-shift assays or [³⁵S]methionine for analysis of the protein products on SDS-polyacrylamide gels. Mobility shift assays were done as described (Evans and Scarpulla 1990). Binding reactions contained 1 μ g of sonicated calf thymus DNA and 2 μ g of BSA in HEPES-D, 100 mM KCl. When indicated, a 200-fold excess of NRF-1-specific or -nonspecific competitors was added. In addition to oligonucleotides described previously (Evans and Scarpulla 1990; Chau et al. 1992), the following oligonucleotides were employed in binding assays:

COXVb -109/-87

GATCCAGAACTGCGCATGTGCGGCGTCA
GTCTTGACGCGTACACGCCGCTGATCGA

mtTFA -78/-49

CGCTCTCCCGCGCCTGCGCCAATT
GGGCGCGGACGCGGTTAAGCGGG

For antiserum supershift experiments, 1 μ l of anti-NRF-1 or preimmune serum was added to binding reactions that had been incubated for 15 min as described above. After an additional 15 min of incubation, the reaction was fractionated by electrophoresis on 4% (58 : 1) acrylamide/bisacrylamide gels.

Expression and purification of the recombinant NRF-1 and antiserum preparation

NRF-1 was expressed using the T7 expression system (Studier et al. 1990). An *NcoI* site was introduced at the NRF-1 initiation codon by PCR using a sense primer, GAACTCCATGGAG-GAACAC, and the same antisense primer as above. The PCR product was digested with *NcoI* and *PstI* to give a 221-bp fragment, ligated with a *PstI*-*BamHI* fragment containing the rest of NRF-1-coding region to the *NcoI* and *BamHI* sites of pET3d and used to transform *E. coli* BL21(DE3). Partial purification of the overexpressed protein was as described (Pognonec et al. 1991). The ammonium sulfate fraction was diluted 10-fold with HEPES-D and applied to a 1-ml heparin-agarose column in HEPES-D, 0.1 M KCl. NRF-1 was eluted in a 0.1-1 M KCl gradient. Goat anti-NRF-1 serum was raised against the heparin-agarose peak fraction (East Acres Biologicals, Southbridge, MA).

Methylation interference and footprinting

Methylation interference and DNase I footprinting were described previously (Evans and Scarpulla 1990). A 130-ng recombinant NRF-1 ammonium sulfate pellet was used in the preparative shift of methylated fragments, and a 20-ng NRF-1 heparin-agarose fraction dialyzed to 0.1 M KCl was used in footprinting. When indicated, a 200-fold excess of NRF-1-specific or -nonspecific competitors was added before the addition of labeled fragment in footprinting.

In vitro transcription

In vitro transcription and analysis of transcripts were done as described (Virbasius et al. 1993), except 54 μ g of HeLa nuclear extract and 0.5 μ g DNA template were used. The recombinant NRF-1 used was the dialyzed heparin-agarose peak fraction. Constructions used as templates have been described (Evans and Scarpulla 1989, 1990).

Transient transfection

For transfection, 3×10^6 - 4×10^6 COS cells were resuspended in 0.8 ml of ZAP buffer (20 mM HEPES, 137 mM NaCl, 0.5 mM KCl, 0.7 mM Na₂HPO₄, 6 mM dextrose adjusted to pH 7.05) and mixed with 5 μ g of reporter plasmid and 15 μ g of pGEM4blue carrier. The cells were then subjected to a single pulse (270 V, 960 μ F) using a Bio-Rad Gene Pulser. The cells were harvested after 48 hr, and extracts were analyzed for chloramphenicol acetyltransferase (CAT) activity and CAT-coding DNA in Hirt supernatants as described previously (Evans and Scarpulla 1988). The reporter plasmids used were either RC4CATB/ Δ -66B or RC4CATB/ Δ -66B with NRF-1 oligonucleotides from MCO5b, mtTFA, or RC4 cloned upstream as described

previously [Evans and Scarpulla 1990]. Values were the average of six separate determinations \pm s.d.

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-1. 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

- Adachi, M. and Y. Tsujimoto. 1990. Potential Z-DNA elements surround the breakpoints of chromosome translocation within the 5' flanking region of bcl-2 gene. *Oncogene* **5**: 1653-1657.
- 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.
- Amy, C.M., B. Williams-Ahlf, J. Naggert, and S. Smith. 1990. Molecular cloning of the mammalian fatty acid synthase gene and identification of the promoter region. *Biochemistry* **27**: 675-679.
- Attardi, G. and G. Schatz. 1988. Biogenesis of mitochondria. *Annu. Rev. Cell Biol.* **4**: 289-333.
- Baer, M., T.W. Nilsen, C. Costigan, and S. Altman. 1990. Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. *Nucleic Acids Res.* **18**: 97-103.
- 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.
- Bender, T.P. and W.M. Kuehl. 1986. Murine myb protooncogene mRNA: cDNA sequence and evidence for 5' heterogeneity. *Proc. Natl. Acad. Sci.* **83**: 3204-3208.
- Biamonti, G., M. Buvoli, M.T. Bassi, C. Morandi, F. Cobianchi, and S. Riva. 1989. Isolation of an active gene encoding human hnRNP protein A1. Evidence for alternative splicing. *J. Mol. Biol.* **207**: 491-503.
- Braidotti, G., I.A. Borthwick, and B.K. May. 1993. Identification of regulatory sequences in the gene for 5-aminolevulinate synthase from rat. *J. Biol. Chem.* **268**: 1109-1117.
- Calzone, F.J., C. Hoog, D.B. Teplow, A.E. Cutting, R.W. Zeller, R.J. Britten, and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo. I. Purification by affinity chromatography and cloning of P3A2, a novel DNA binding protein. *Development* **112**: 335-350.
- Cavaillès, V., P. Augereau, and H. Rochefort. 1993. Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells. *Proc. Natl. Acad. Sci.* **90**: 203-207.
- Chang, D.D. and D.A. Clayton. 1989. Mouse RNase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. *Cell* **56**: 131-139.
- 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 2 α , 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.
- Das, A.T., A.C. Arnberg, H. Malingre, P. Moerer, R. Charles, A.F. Moorman, and W.H. Lamers. 1993. Isolation and characterization of the rat gene encoding glutamate dehydrogenase. *Eur. J. Biochem.* **211**: 795-803.
- Degen, S.J., L.A. Stuart, S. Han, and C.S. Jamison. 1991. Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor-like protein: Expression during development. *Biochemistry* **30**: 9781-9791.
- Desimone, S.M. and K. White. 1993. The *Drosophila* erect wing gene, which is important for both neuronal and muscle development, encodes a protein which is similar to the sea urchin P3A2 DNA binding protein. *Mol. Cell. Biol.* **13**: (6) 3641-3949.
- Diffley, J.F. and B. Stillman. 1991. A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc. Natl. Acad. Sci.* **88**: 7864-7868.
- 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.
- Fisher, R.P., J.N. Topper, and D.A. Clayton. 1987. Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell* **50**: 247-258.
- Forrest, G.L., S. Akman, J. Doroshov, H. Rivera, and W.D. Kaplan. 1991. Genomic sequence and expression of a cloned human carbonyl reductase gene with daunorubicin reductase activity. *Mol. Pharmacol.* **40**: 502-507.
- Furuno, N., K. Nakagawa, U. Eguchi, M. Ohtsubo, T. Nishimoto, E. Soeda, and M. Ohtsubo. 1991. Complete nucleotide sequence of the human RCC1 gene involved in coupling between DNA replication and mitosis. *Genomics* **11**: 459-461.
- Giallongo, A., D. Oliva, L. Cali, G. Barba, G. Barbieri, and S. Feo. 1990. Structure of the human gene for alpha-enolase. *Eur. J. Biochem.* **190**: 567-573.
- Green, S., I. Isseman, and E. Sheer. 1988. A versatile *in vivo* and *in vitro* eukaryotic expression vector for protein engineering. *Nucleic Acids Res.* **16**: 369.
- Haendler, B. and E. Hofer. 1990. Characterization of the human cyclophilin gene and of related processed pseudogenes. *Eur. J. Biochem.* **190**: 477-482.
- Hoog, C., F.J. Calzone, A.E. Cutting, R.J. Britten, and E.H. Davidson. 1991. Gene regulatory factors of the sea urchin embryo. II. Two dissimilar proteins, P3A1 and P3A2, bind to the same target sites that are required for early territorial gene expression. *Development* **112**: 351-364.
- Hurt, M.M., N. Chodchoy, and W.F. Marzluff. 1989. The mouse histone H2a.2 gene from chromosome 3. *Nucleic Acids Res.* **17**: 8876.
- Jenkins, E.P., C.L. Hsieh, A. Milatovich, K. Normington, D.M. Berman, U. Francke, and D.W. Russell. 1991. Characterization and chromosomal mapping of a human steroid 5 alpha-

- reductase gene and pseudogene and mapping of the mouse homologue. *Genomics* **11**: 1102–1112.
- Kadonaga, J.T. 1991. Purification of sequence-specific binding proteins by DNA affinity chromatography. *Methods Enzymol.* **208**: 10–12.
- Katz, A. and C. Kahana. 1988. Isolation and characterization of the mouse ornithine decarboxylase gene. *J. Biol. Chem.* **263**: 7604–7609.
- Khew-Goodall, Y., R.E. Mayer, F. Maurer, S.R. Stone, and B.A. Hemmings. 1991. Structure and transcriptional regulation of protein phosphatase 2A catalytic subunit genes. *Biochemistry* **30**: 89–97.
- Kreysing, J., K. von Figura, and V. Gieselmann. 1990. Structure of the arylsulfatase A gene. *Eur. J. Biochem.* **191**: 627–631.
- Ku, D.H., S.C. Wen, A. Engelhard, N.C. Nicolaides, K.E. Lipson, T.A. Marino, and B. Calabretta. 1993. c-myc transactivates cdc2 expression via Myb binding sites in the 5'-flanking region of the human cdc2 gene. *J. Biol. Chem.* **268**: 2255–2259.
- Kunze, N., G.C. Yang, M. Dolberg, R. Sundarp, R. Knippers, and A. Richter. 1991. Structure of the human type I DNA topoisomerase gene. *J. Biol. Chem.* **266**: 9610–9616.
- Kurihara, T., K. Monoh, K. Sakimura, and Y. Takahashi. 1990. Alternative splicing of mouse brain 2',3'-cyclic-nucleotide 3'-phosphodiesterase mRNA. *Biochem. Biophys. Res. Commun.* **170**: 1074–1081.
- Kutt, H., J. Herz, and K.K. Stanley. 1989. Structure of the low-density lipoprotein receptor-related protein (LRP) promoter. *Biochim. Biophys. Acta* **1009**: 229–236.
- Lee, C.-M., R.S. Haun, S.-C. Tsai, J. Moss, and M. Vaughan. 1992. Characterization of the human gene encoding ADP-ribosylation factor 1, a guanine nucleotide-binding activator of cholera toxin. *J. Biol. Chem.* **267**: 9028–9034.
- Luethy, J.D., J. Fagnoli, J.S. Park, A.J. Fornace, Jr., and N.J. Holbrook. 1990. Isolation and characterization of the hamster gadd153 gene. Activation of promoter activity by agents that damage DNA. *J. Biol. Chem.* **265**: 16521–16526.
- Lutfalla, G., K. Gardiner, D. Proudhon, E. Vielh, and G. Uze. 1992. The structure of the human interferon alpha/beta receptor gene. *J. Biol. Chem.* **267**: 2802–2809.
- Mamula, P.W. and I.D. Goldfine. 1992. Cloning and characterization of the human insulin-like growth factor-I receptor gene 5'-flanking region. *DNA Cell Biol.* **11**: 43–50.
- Mitsubuchi, H., Y. Nobukuni, F. Endo, and I. Matsuda. 1991. Structural organization and chromosomal localization of the gene for the E1 beta subunit of human branched chain alpha-keto acid dehydrogenase. *J. Biol. Chem.* **266**: 14686–14691.
- Miyake, S., Y. Emori, and K. Suzuki. 1986. Gene organization of the small subunit of the human calcium-activated neutral protease. *Nucleic Acids Res.* **14**: 8805–8807.
- Miyatake, S., T. Otsuka, T. Yokota, F. Lee, and K. Arai. 1985. Structure of the chromosomal gene for granulocyte-macrophage colony stimulating factor: Comparison of the mouse and human genes. *EMBO J.* **4**: 2561–2568.
- Moshier, J.A., D.L. Osborne, M. Skunca, J. Dosescu, J.D. Gilbert, M.C. Fitzgerald, G. Polidori, R.L. Wagner, R. Friezner, S.J. Degen, and G.D. Luk. 1992. Multiple promoter elements govern expression of the human ornithine decarboxylase gene in colon carcinoma cells. *Nucleic Acids Res.* **20**: 2581–2590.
- Nagley, P. 1991. Coordination of gene expression in the formation of mammalian mitochondria. *Trends Genet.* **7**: 1–4.
- Nojima, H. 1989. Structural organization of multiple rat calmodulin genes. *J. Mol. Biol.* **208**: 269–282.
- 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.
- Park, J.S., J.D. Luethy, M.G. Wang, J. Fagnoli, A.J. Fornace, Jr., O.W. McBride, and N.J. Holbrook. 1992. Isolation, characterization and chromosomal localization of the human GADD153 gene. *Gene* **116**: 259–267.
- Parmentier, M. and A. Lefort. 1991. Structure of the human brain calcium-binding protein calretinin and its expression in bacteria. *Eur. J. Biochem.* **196**: 79–85.
- Pearson, B.E., H.P. Nasheuer, and T.S. Wang. 1991. Human DNA polymerase alpha gene: Sequences controlling expression in cycling and serum-stimulated cells. *Mol. Cell. Biol.* **11**: 2081–2095.
- 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.
- Sauerwald, A., C. Hoesche, R. Oschwald, and M.W. Kilmann. 1990. The 5'-flanking region of the synapsin I gene. A G + C-rich, TATA- and CAAT-less, phylogenetically conserved sequence with cell type-specific promoter function. *J. Biol. Chem.* **265**: 14932–14937.
- Sittman, D.B., R.A. Graves, and W.F. Marzluff. 1983. Structure of a cluster of mouse histone genes. *Nucleic Acids Res.* **11**: 6679–6697.
- 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.
- Tasanen, K., J. Oikarinen, K.I. Kivirikko, and T. Pihlajaniemi. 1992. Promoter of the gene for the multifunctional protein disulfide isomerase polypeptide. Functional significance of the six CCAAT boxes and other promoter elements. *J. Biol. Chem.* **267**: 11513–11519.
- Tewari, D.S., D.M. Cook, and R. Taub. 1989. Characterization of the promoter region and 3' end of the human insulin receptor gene. *J. Biol. Chem.* **264**: 16238–16245.
- Tominaga, K., S. Akiyama, Y. Kagawa, and S. Ohta. 1992. Upstream region of a genomic gene for human mitochondrial transcription factor 1. *Biochim. Biophys. Acta* **1131**: 217–219.
- Topper, J.N. and D.A. Clayton. 1990. Characterization of human MRP/Th RNA and its nuclear gene: Full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res.* **18**: 793–799.
- Tsukamoto, T., R. Toyama, H. Itoh, T. Kozasa, M. Matsuoka, and Y. Kaziro. 1991. Structure of the human gene and two rat cDNAs encoding the alpha chain of GTP-binding regulatory protein Go: Two different mRNAs are generated by alternative splicing. *Proc. Natl. Acad. Sci.* **88**: 2974–2978.
- 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.
- . 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.* (in press).
- 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.
- Wagner, M. and R.P. Perry. 1985. Characterization of the multi-gene family encoding the mouse S16 ribosomal protein: Strategy for distinguishing an expressed gene from its processed pseudogene counterparts by an analysis of total genomic DNA. *Mol. Cell. Biol.* **5**: 3560–3576.

- Wallace, D.C. 1992. Diseases of the mitochondrial DNA. *Annu. Rev. Biochem.* **61**: 1175–1212.
- . 1993. Mitochondrial diseases: Genotype versus phenotype. *Trends Genet.* **9**: 128–133.
- Walter, T., B. Drabent, H. Krebs, M. Tomalak, S. Heiss, and B.J. Benecke. 1989. Cloning and analysis of a human 86-kDa heat-shock-protein-encoding gene. *Gene* **83**: 105–115.
- Wen, L., J.K. Huang, and P.J. Blackshear. 1989. Rat ornithine decarboxylase gene. Nucleotide sequence, potential regulatory elements, and comparison to the mouse gene. *J. Biol. Chem.* **264**: 9016–9021.
- Xu, B. and D.A. Clayton. 1992. Assignment of a yeast protein necessary for mitochondrial transcription initiation. *Nucleic Acids Res.* **20**: 1053–1059.
- Yagawa, Y., K. Kawakami, and K. Nagano. 1990. Cloning and analysis of the 5'-flanking region of rat Na⁺/K⁺-ATPase alpha 1 subunit gene. *Biochim. Biophys. Acta* **1049**: 286–292.
- Zhou, Q.Y., C. Li, and O. Civelli. 1992. Characterization of gene organization and promoter region of the rat dopamine D1 receptor gene. *J. Neurochem.* **59**: 1875–1883.



NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators.

C A Virbasius, J V Virbasius and R C Scarpulla

Genes Dev. 1993, **7**:

Access the most recent version at doi:[10.1101/gad.7.12a.2431](https://doi.org/10.1101/gad.7.12a.2431)

References

This article cites 71 articles, 31 of which can be accessed free at:
<http://genesdev.cshlp.org/content/7/12a/2431.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](#).

