Molecular Cloning and Characterization of a Novel Mouse Macrophage Gene That Encodes a Nuclear Protein Comprising Polyglutamine Repeats and Interspersing Histidines*

(Received for publication, January 25, 1996, and in revised form, June 18, 1996)

George W. Cox‡§, Lynn S. Taylor‡, Jonathan D. Willis‡§, Giovanni Melillo‡, Robert L. White III‡¶, Stephen K. Anderson¶, and Jih-Jing Lin**

From the Laboratories of Experimental Immunology and Biochemical Physiology, Division of Basic Sciences and Biological Carcinogenesis and Development Program, Science Applications International Corporation (SAIC) Frederick, National Cancer Institute-Frederick Cancer Research and Development Center, National Institutes of Health, Frederick, Maryland 21702-1201

Simple tandem repeats of the trinucleotide sequence CAG encode homopolymeric stretches of glutamine. Although polyglutamine has been identified in diverse proteins, it is present predominantly in transcription factors. We observed that oncogene-immortalized mouse macrophages express several genes that contain a CAG repeat motif. Therefore, we attempted to clone a novel gene that contains a CAG repeat and is associated with cytokine activation of macrophages. Screening of a mouse macrophage cDNA library with a probe comprising 12 consecutive CAG triplets identified at least one unique clone. The cDNA encodes a protein (named GRP-1 or glutamine repeat protein-1) with 171 amino acids, a calculated molecular mass of 21.6 kDa, and a predicted pl of 10.67. Greater than two-thirds of GRP-1 are only two amino acids, namely glutamine (50%) and histidine (18%). There are four polyglutamine motifs interspersed with histidine-rich regions. There is also a putative nuclear localization signal flanked by sites for possible serine phosphorylation. GRP-1 mRNA was expressed constitutively in some macrophage cell lines and B and T cell lines. Interferon-γ or lipopolysaccharide augmented GRP-1 mRNA expression in the mouse macrophage cell line ANA-1. Western blot analyses using an antipeptide serum revealed that GRP-1 was localized in the nucleus of ANA-1 macrophages and transfected 3T3 fibroblasts. Overexpression of GRP-1 decreased Sp1-driven chlormehanol acetyltransferase gene expression in transient cotransfection experiments. Because polyglutamine motifs can cause protein oligomerization and can function as transcriptional activation domains, we suggest that GRP-1 may be a transcription factor associated with interferon-γ- or lipopolysaccharide-induced activation of macrophages.

Macrophages play a major role in the immune system through numerous activities expressed either constitutively or after exposure to cytokines, bacterial lipopolysaccharide (LPS)1 or other agents either alone or in combination (1). For example, macrophages affect T-cell activation by processing and presenting antigens, and they promote inflammatory reactions by secreting various cytokines. In addition, activated macrophages exert potent antibacterial, antiviral, and antitumor functions.

To avoid the difficulties associated with isolating large homogeneous populations of macrophages for studying such functions, several oncogene-immortalized mouse macrophage cell lines were generated that emulate normal tissue macrophages (2–4). Although these and other macrophage populations do not synthesize and secrete IL-2, Northern blots analyses revealed the expression of several genes in oncogene-immortalized macrophages that were at least partially homologous with mouse IL-2 cDNA. The expression of at least one of these genes was regulated by IFN-γ and/or LPS. Subsequent experiments showed that these macrophage genes were homologous with sequences contained in exon I of the mouse IL-2 gene. This region is particularly noteworthy because it contains a tandem repeat of the trinucleotide sequence CAG (5, 6).

Simple tandem repeats of the trinucleotide sequence CAG encode tracts of glutamine and are present in many diverse genes. Wharton et al. (7) identified CAG repeats in the notch locus of Drosophila melanogaster and named these sequences opa. Similar repeats were also identified in many other developmentally regulated loci or transcription factor genes of D. melanogaster (7, 8), and opa transcripts were molecularly cloned in the mouse (9). Many mammalian DNA-binding transcription factors, including the androgen receptor (10), the glucocorticoid receptor (11, 12), c-myc (13, 14), the TATA-binding protein of transcription factor IID (15), and Sp1 (16), contain tracts of glutamine or glutamine-rich domains. Indeed, polyglutamine motifs are present in more than 30 transcription factors. Dynamic mutation of simple tandem repeats (including those encoding polyglutamine) has been implicated recently in the molecular pathogenesis of several human genetic diseases (17, 18). The onset and severity of some inherited neurodegenerative diseases (e.g. Huntington’s disease) have been attributed to expansion of CAG repeats and polyglutamine (19, 20).

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank‡¶/EBI Data Bank with accession number(s) U46463.
¶ To whom correspondence should be addressed: Dept. of Pharmacology, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Rd., C2021, Bethesda, MD 20814-4799. Tel.: 301-295-3232; Fax: 301-295-3229.
§ Participant in the Werner H. Kirsten Student Intern Program and the Student Research Training Program of the National Cancer Institute, National Institutes of Health.

† The abbreviations used are: LPS, lipopolysaccharide; IL, interleukin; IFN-γ, interferon-γ; bp, base pair(s); GRP-1, glutamine repeat protein-1; CAT, chlormehanol acetyltransferase; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; pBS, pBlue-script SK5- plasmid; pBS-GRP-1, pBS containing bp 129-1968 of GRP-1 cDNA; DMEM, Dulbecco’s modified Eagle’s medium; NLS, nuclear localization signal.
** G. W. Cox, L. S. Taylor, J. D. Willis, G. Melillo, R. L. White III, S. K. Anderson, and J.-J. Lin, unpublished observations.
In spinal and bulbar muscular atrophy (also known as Kennedy disease), the affected protein is a known transcription factor, namely the androgen receptor (21, 22).

The functions of trinucleotide repeats and glutamine repeats have not been determined fully. It has been suggested that these motifs may form spacer or hinge regions between functional areas (23), cause the formation of hairpin structures (24), and result in the alteration of chromatin structure (25). Perutz et al. (26, 27) have suggested that glutamine repeats can form polar zippers. A recent study confirmed that glutamine repeats can cause protein oligomerization (28) and supported accumulating evidence that polyglutamine can act as a protein interaction motif. Kadonaga et al. (29), Courey et al. (30), and Courey and Tjian (31) examined the function of a glutamine-rich region in transcription factor Sp1 and showed that it is an activation domain. In addition, Seipel et al. (32) and Gerber et al. (33) demonstrated that polyglutamine can stimulate the transcription of particular reporter genes when it is fused to the DNA binding domain of the yeast protein GAL4. In contrast, polyglutamine can inhibit transcription, and its expansion in the androgen receptor decreases transactivation function (22, 34).

Cytokines and LPS affect many intracellular events, including regulation of gene transcription, shown to be involved in or as a consequence of macrophage activation. Oncogene-immortalized macrophage cell lines may express novel IFN-γ and/or LPS-responsive genes that contain repeats of the trinucleotide CAG. Therefore, we attempted to clone a novel gene associated with cytokine and/or LPS-induced activation of macrophages that contains a CAG repeat and may encode a transcription factor. We report here the molecular cloning, expression, and characterization of a novel gene that encodes a nuclear protein (named GRP-1 or glutamine repeat protein-1) comprising several polyglutamine tracts and interspersing histidines. GRP-1 mRNA was expressed constitutively in macrophages and other cells, and its expression was augmented in cytokine- or LPS-treated macrophages. Preliminary analysis of the function of GRP-1 demonstrated that its overexpression impairs Sp1-driven chloramphenicol acetyltransferase (CAT) gene expression in transiently transfected 3T3 fibroblasts.

MATERIALS AND METHODS

Cells—The mouse macrophage cell lines GG2EE and ANA-1 were generated previously by infecting fresh bone marrow-derived cells from C3H/HeJ or C57BL/6 mice, respectively, with the murine recombinant retrovirus J2 (expressing the v-rel and v-myc oncoproteins) (3, 4). Animal care was provided in accordance with the procedures outlined in the “Guide for the Care and Use of Laboratory Animals” (NIH Publication No. 86-23, 1985). Peritoneal exudate macrophages were harvested from C57BL/6 mice 4 days after intraperitoneal injection of 1 ml of 3% Brewer thioglycollate medium (BBL Microbiology Systems, Becton Dickinson, Cockeysville, MD). The cell lines P388D1, CTLL-2, PC-12, and P815 were obtained from American Type Culture Collection (Rockville, MD). The cell line RL(male)/1 was provided by Dr. L. Mason (National Cancer Institute, Frederick, MD). The cell lines YAC-1 and EL-4 were gifts from Dr. Q. Hu (National Cancer Institute, Frederick, MD). The cell line 3T3 (BALB/c origin) was from Dr. M. Smith (SAIC-Frederick, Frederick, MD). The macrophage cell lines and 3T3 fibroblasts were routinely cultured in DMEM (Whittaker Bioproducts, Walkersville, MD) supplemented with 10% heat-inactivated fetal bovine serum (HyClone Laboratories, Logan, UT), 2 mmol/l glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin (Life Technologies, Inc.). Other cells were cultured in either DMEM or RPMI 1640 supplemented as above. Cell cultures were maintained at 37 °C in a humidified incubator containing 5% CO2 in air.

cDNA Library Construction and Screening—The LPS-hyperresponsive macrophage cell line GG2EE was treated for approximately 18 h with 100 units/ml of IFN-γ (Life Technologies, Inc.) plus 18,000 IU/ml of highly purified human recombinant IL-2 (derived from Escherichia coli) (Chemicon International, CA). Poly(A)+ RNA was extracted from the cells using the FastTrack mRNA isolation kit (Invitrogen Corp., San Diego, CA). A custom cDNA library was constructed by Stratagene Cloning Systems (La Jolla, CA). Briefly, oligo(dT)-primed cDNA fragments of >500 bp were cloned in the EcoR1XhoI site of Uni-ZAP XR vector. The number of primary plaques was 8.8 × 106, and the estimated background was <2% nonrecombinant. The estimated unamplified template was 2.2 × 108 plaque-forming units/ml. Approximately 106 A phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA using standard methods. Nucleic acid radiolabeling and hybridization were performed essentially as described below. Positive clones (~30) were subjected to two additional screenings. The final screen resulted in 17 positive phage clones subject to an antisense to an in vivo excision protocol (according to B. Stamatoyannopoulos) to rescue pBS containing the cloned cDNA insert. One clone (ZAP-E clone 9) was partially sequenced using the dyeoxy chain termination method (37) with Sequenase version 2.0 T7 DNA polymerase (U. S. Biochemical Corp.). The cDNA library was rescreened (~500,000 plaques) with ZAP-E clone 9 cDNA to obtain an independent clone (CL10A, also known as GRP-1 cDNA), which was subsequently isolated in pBS. The complete sequence of the sense strand and the complete sequence of the antisense strand of GRP-1 cDNA in pBS were obtained using overlapping oligonucleotide primers. The 5′ end of GRP-1 cDNA was obtained by using the 5′-RACE system for rapid amplification of cDNA ends (Life Technologies, Inc.). The 5′-RACE product was purified, appropriately inserted into pBS-GRP-1, and sequenced as described above. All sequence analyses were performed using the Sequencher Analysis Package by Genetics Computer Group, Inc. (Madison, WI).

Southern Blot Analysis—Pure high molecular weight genomic DNA was isolated from ANA-1 macrophages using the QIAgen Cell Culture DNA kit (QIAGEN, Chatsworth, CA). Ten μg of DNA were digested to completion with EcoRI, BamHI, or HindIII (Life Technologies, Inc.). After extraction with phenol/chloroform and ethanol precipitation, the digested DNAs were electrohoresed in 0.8% agarose. The DNAs in the agarose gel were subjected to denaturation and partial hydrolysis by acid depurination. The DNAs were transferred to Nitray membrane (Schleicher & Schuell) using 10× SSC and were cross-linked to the membrane by ultraviolet irradiation. Prehybridization was performed with HybriZol (Boehringer Mannheim). After hybridization with HybriZol I or hybrid with HybriZol I overnight at 42 °C, the blot was washed three times in 2× SSC/0.1% SDS for 5 min at room temperature and then washed twice in 0.1× SSC/0.1% SDS for 15 min at 60 °C. The blot was autoradiographed at −70 °C on XAR-5 film (Eastman Kodak Co.) with the use of Lightning Plus intensifying screens (DuPont NEN).

RNA Preparation and Reverse Transcription-PCR—Total cellular RNAs were purified using TriZOL RNA isolation reagent (39) (Life Technologies, Inc.). First-strand cDNA synthesis and PCR were performed essentially as described previously (40). One μg of RNA in water was denatured for 5 min at 90 °C, quick chilled on ice, and incubated with the following reagents at the indicated final concentrations in a total volume of 20 μl for 1 h at 42 °C: dCTP (1 mm), dATP (1 mm), dGTP (1 mm), dTTP (1 mm) (all from Perkin-Elmer), RNASin (1 unit/μl) (Promega Corp., Madison, WI), PCR reaction buffer (1×) (Perkin-Elmer), random hexamer oligonucleotide (5 pmol/μl) (Pharmacia Biotech Inc.), and SuperScript II RNase H− reverse transcriptase (10 units/μl) (Life Technologies). The reaction was terminated by heating for 5 min at 95 °C and quick chilled on ice. The entire 20-μl reaction was subjected to PCR with the following reagents at the indicated final concentrations in a total volume of 100 μl: 0.5 pmol/μl GRP-1 cDNA sense primer (5′-TGGTGGCATATTGACTGT-3′), 0.5 pmol/μl GRP-1 cDNA antisense primer (5′-TCCACCCCATCTTGGTGT-3′), 0.2 μl/μl Taq polymerase (50 units/μl) (Perkin-Elmer), and PCR reaction buffer (1×). PCR was performed using the step-cycle program on a DNA Thermal Cycler (Perkin-Elmer) as follows: 95 °C (1 min), 55 °C (2 min), 72 °C (3 min) for 30 cycles, and then 4 °C (soak). Approximately one-third of the PCR product was Southern blotted and probed with [32P]dCTP-labeled GRP-1 cDNA using standard methods (ZAP-E clone 9).

In Vitro Transcription and Translation—pBS-GRP-1 was subjected to in vitro transcription/translation using [35S]methionine (Amersham Corp.), T3 RNA polymerase (Promega Corp.), and the TNT Coupled Reticulocyte Lysate System (Promega Corp.). The translation products were separated by SDS-PAGE under reducing conditions (41) and autoradiographed.
Cloning and Expression of Glutamine Repeat Protein-1

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 (i.e., Met) is present at amino acids 1–3 (47). The nucleotide sequence is presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

by thin-layer chromatography (46). In addition to using cotransfection of pGL3-control plasmid followed by luciferase assay as a control for differences in transfection efficiency, we prepared genomic DNA from replicates and performed Southern blot analysis using pGL3-control plasmid as a probe.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

by thin-layer chromatography (46). In addition to using cotransfection of pGL3-control plasmid followed by luciferase assay as a control for differences in transfection efficiency, we prepared genomic DNA from replicates and performed Southern blot analysis using pGL3-control plasmid as a probe.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

by thin-layer chromatography (46). In addition to using cotransfection of pGL3-control plasmid followed by luciferase assay as a control for differences in transfection efficiency, we prepared genomic DNA from replicates and performed Southern blot analysis using pGL3-control plasmid as a probe.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

by thin-layer chromatography (46). In addition to using cotransfection of pGL3-control plasmid followed by luciferase assay as a control for differences in transfection efficiency, we prepared genomic DNA from replicates and performed Southern blot analysis using pGL3-control plasmid as a probe.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.

by thin-layer chromatography (46). In addition to using cotransfection of pGL3-control plasmid followed by luciferase assay as a control for differences in transfection efficiency, we prepared genomic DNA from replicates and performed Southern blot analysis using pGL3-control plasmid as a probe.

RESULTS

Cloning and Sequencing of GRP-1 cDNA—Poly(A)⁺ RNA was isolated from IFN-γ plus IL-2-treated GG2EE mouse macrophages and was used to construct a cDNA library in the Uni-ZAP XR vector. Approximately 10⁶ phage plaques were screened with a radiolabeled probe corresponding to bp 108–236 of mouse IL-2 cDNA. Several positive clones were examined in Northern blot analyses for their ability to hybridize with mRNA isolated from unstimulated or stimulated ANA-1 mouse macrophages (data not shown). GRP-1 cDNA was initially isolated as a 1840-bp clone that detected two mRNA transcripts of approximately 2.0 and 3.5 kb in ANA-1 macrophages. The 5'-RACE procedure was used to obtain the nearly full-length 1968-bp GRP-1 cDNA clone. The nucleotide sequence and deduced amino acid sequence of GRP-1 are shown in Fig. 1. An ATG triplet beginning at nucleotide position 181 corresponds to the predicted initiation codon. The putative NLS is double underlined. The putative mRNA instability sequences are presented in italicized lowercase letters. The polyadenylation signal sequence is underlined.
sites for protein kinase C-mediated (48, 49) and casein kinase II-mediated (50) phosphorylation, respectively. There are no potential N-linked glycosylation sites present in the predicted protein. The 3'-untranslated region of the gene contains four copies of the mRNA instability consensus sequence ATTTA (51). The prototypical polyadenylation signal AATAAA (52) is not present. However, the variant signal ATTAAA (53, 54) begins 20-bp upstream of the poly(A) tail.

Fig. 2 depicts the results of a Kyte-Doolittle algorithm analysis (55) of the deduced amino acid sequence of GRP-1. The presence of a large span of hydrophilic residues indicates that the protein is extremely polar. Although there are two hydrophobic regions, present at the N and C termini, neither one is long enough to constitute an N terminus signal peptide nor a membrane spanning domain typical of secreted or transmembrane proteins. Collectively, these data suggest that GRP-1 is intracellular.

Data Base Searches—A search of the current nonredundant combination of all nucleic acid and protein data bases with GRP-1 nucleotide and amino acid sequences revealed homology to many other sequences. However, significant homology was limited to the tandem CAG repeats (7, 9, 56) and polyglutamine tracts common among homeobox proteins (e.g. notch, zeste, mastermind, and fsh) and other transcription factors (e.g. the glucocorticoid and androgen receptors, c-myc and the TATA-binding protein of transcription factor IID).

Southern Blot Analysis—Genomic DNA from ANA-1 macrophages was digested with EcoRI, BamHI, or HindIII and was subjected to electrophoresis and blotting. The blot was incubated with a radiolabeled probe without CAG repeats. As shown in Fig. 3, the cDNA probe hybridized with a single band of genomic DNA from each digestion.

Expression of GRP-1 mRNA in Various Cells—To begin to examine the expression of the GRP-1 gene in various cells, total RNA was isolated from several types of cells and was subjected to reverse transcription-PCR analysis using GRP-1 gene-specific primers. The cDNA products were Southern blotted and probed with radiolabeled cDNA corresponding to the expected PCR product. The data show that GRP-1 mRNA was expressed constitutively in oncogene-immortalized macrophages (ANA-1 and GG2EE cell lines) and normal thiglycollate-elicited peritoneal exudate macrophages (Fig. 4). In addition, several T-cell populations including the lymphomas YAC-1 and RL(male)1 and normal fetal thymus cells expressed GRP-1 mRNA. Although the IL-2-dependent cytotoxic T-cell line CTL-2 expressed GRP-1 mRNA constitutively, the T-cell lymphoma (thymoma) EL-4 that can secrete IL-2 did not express detectable basal levels of GRP-1 mRNA. The pre-B cell line 70Z/3, but not the more mature B-cell leukemia/lymphoma BCL1, also expressed GRP-1 mRNA constitutively. At least one cell line of nonhematopoietic origin, i.e. the renal carcinoma Renca, appeared to express relatively high constitutive levels of GRP-1 mRNA. The myeloid cell lines P388D1 and M1, the mastocytoma cell line F515, and the rat pheochromocytoma cell line PC-12 did not express detectable levels of GRP-1 mRNA. Overall, these data show that GRP-1 mRNA is expressed constitutively in several different types of cells.

Translation of GRP-1 mRNA in Vitro—To verify that the cloned GRP-1 cDNA represents a transcript that can be translated, we performed in vitro transcription and translation analyses. GRP-1 cDNA contained in pBS was transcribed in vitro using T3 RNA polymerase, and the resultant mRNA was translated in vitro using rabbit reticulocyte lysate. As shown by the SDS-PAGE analysis depicted in Fig. 5, lane 1, undigested pBS-GRP-1 generated two prominent bands of protein with molecular weights of 30,000–35,000. To demonstrate that the detected proteins were encoded by sequences contained in the predicted open reading frame, pBS-GRP-1 was pretreated with either SspI that digests GRP-1 cDNA at nucleotide position 720 or AvaII that digests GRP-1 cDNA at nucleotide position 306. The AvaII-digested pBS-GRP-1 failed to generate detectable bands of protein (Fig. 5, lane 3). However, the SspI-digested pBS-GRP-1 generated bands of protein that were similar to those generated by the undigested pBS-GRP-1 (Fig. 5, lane 2). No bands of protein were detected in reactions lacking pBS-GRP-1 (Fig. 5, lane 4).

Overexpression of GRP-1 in 3T3 Fibroblasts—To examine whether the cloned GRP-1 cDNA can be transcribed and translated in intact cells, 3T3 fibroblasts were transfected with pcDNA3 (expressing neomycin resistance alone) or with pcDNA3 carrying GRP-1 cDNA in the sense or antisense ori-
Western blot analysis using a rabbit antiserum raised against a synthetic GRP-1 C terminus peptide failed to detect any bands of protein in either the cytosolic fraction or the nuclear extract of 3T3 fibroblasts transfected with an empty expression vector (Fig. 6) or expression vector carrying GRP-1 cDNA in the antisense orientation (data not shown). However, a predominant band of protein with a molecular weight of approximately 33,000 was detected in both the cytosolic fraction and nuclear extract of 3T3 fibroblasts transfected with expression vector carrying GRP-1 cDNA in the sense orientation.

Expression of GRP-1 in ANA-1 Macrophages—The next experiment was performed to examine the expression and subcellular localization of GRP-1 in ANA-1 macrophages. Cytosol and nuclear extract were prepared from untreated ANA-1 macrophages and were subjected to Western blot analysis as described above. The results show that GRP-1 was expressed constitutively and was localized predominantly in the nucleus of ANA-1 macrophages (Fig. 7). The data also confirm the specificity of the antiserum because preincubation of the antiserum with unconjugated synthetic peptide prevented subsequent detection of GRP-1.

Regulation of GRP-1 Gene Expression in ANA-1 Macrophages—Experiments were performed to determine whether prototypical activators of macrophages regulate the expression of GRP-1 mRNA in ANA-1 macrophages. As shown in Fig. 8, ANA-1 macrophages expressed low levels of GRP-1 mRNA constitutively. However, IFN-γ alone, LPS alone, or IL-2 alone augmented GRP-1 mRNA expression.

Overexpression of GRP-1 Decreases Sp1-driven CAT Expression—As discussed above, polyglutamine causes protein oligomerization (28) and likely acts as a general protein interaction motif in transcription factors (57) and other proteins. Based on its nuclear localization and abundance of glutamines/polyl glutamines, we hypothesize that GRP-1 is a transcription factor. To begin to examine its ability to regulate transcription via glutamine-dependent protein-protein interactions, we tested whether GRP-1 can affect the activity of a transcription factor that contains glutamine-rich activation (i.e. interaction) domains. As shown in Fig. 9, upper panel, increasing amounts

Fig. 5. Translation of GRP-1 mRNA in vitro. The nearly full-length GRP-1 cDNA that was cloned in the EcoRI/XhoI site of pBS was transcribed into mRNA in the sense orientation using T3 polymerase. The mRNA was subsequently translated using rabbit reticulocyte lysate in the presence of [35S]methionine as described under “Materials and Methods.” Lane 1 shows the results obtained with undigested plasmid DNA. Lane 2 shows the results obtained with SspI-digested plasmid DNA. Lane 3 shows the results obtained with AvaII-digested plasmid DNA. Lane 4 shows the results obtained without plasmid DNA.

Fig. 6. Overexpression of GRP-1 in 3T3 fibroblasts. 3T3 fibroblasts were transiently transfected with pcDNA3 alone (neo, neomycin resistance) or pcDNA3 carrying GRP-1 cDNA in the sense orientation (GRP-1). Twenty μg of cytosolic protein or nuclear extract were subjected to Western blot analysis using a rabbit antiserum raised against a synthetic peptide corresponding to the last 13 amino acids of GRP-1 as described under “Materials and Methods.”

Fig. 7. Expression and subcellular localization of GRP-1 in ANA-1 macrophages. Twenty μg of cytosolic protein or nuclear extract from unstimulated ANA-1 macrophages were subjected to Western blot analysis using an antipeptide serum as described above and under “Materials and Methods.” For peptide competition analysis, 25–50 μg of unconjugated synthetic peptide were incubated with antiserum for 1 h before incubation with the blot as described under “Materials and Methods.”
of pcDNA3/GRP-1 caused a dose-dependent decrease (up to approximately 60%) of Sp1-driven CAT expression in transiently cotransfected 3T3 fibroblasts. Overexpression of GRP-1 also decreased SV40-driven luciferase activity, precluding our ability to use cotransfection of pGL3-control plasmid as a control for equal DNA uptake among the various samples (Fig. 9, lower panel). However, we judged the transfection efficiencies to be equivalent based on Southern blot analysis (data not shown).

DISCUSSION

We have described the molecular cloning of a novel gene characterized primarily by an abundance of CAG or opa (7) repeats. The encoded protein, GRP-1, is remarkable because it contains mostly glutamines (50%) and histidines (18%). These amino acids are not distributed randomly but appear as tracts of polyglutamine interspersed with histidine-glutamine-glutamine reiterations. GRP-1 mRNA was expressed constitutively in several cell types, and IFN-γ or LPS augmented its expression in ANA-1 macrophages. GRP-1 was localized in the nucleus of ANA-1 macrophages and transient 3T3 fibroblasts. Overexpression of GRP-1 decreased Sp1-driven CAT expression and SV40-driven luciferase expression in cotransfection experiments. Because polyglutamine and glutamine-rich motifs function as activation domains in many transcription factors, we hypothesize that GRP-1 is a transcription factor associated with the IFN-γ- or LPS-induced activation of macrophages.

Sequencing of GRP-1 cDNA revealed an ATG triplet beginning at nucleotide position 39 that conforms extremely well to the optimal context for initiation of translation (GCCA/GC-GAUGG) described by Kozak (58, 59). However, it is followed shortly afterwards by an in-frame stop codon (UGA) beginning at nucleotide position 135. Therefore, we assumed that the ATG beginning at nucleotide position 181 corresponds to the initiation codon. Although it conforms less well to the Kozak consensus sequence (except the presence of a C at −1, a G at −6, and an A at +4), it is preceded by an in-frame stop codon and generates the longest possible open reading frame (i.e. 513 bp). The deduced amino acid sequence predicts a very unusual protein with a disproportionate composition of glutamines and histidines. The histidine-glutamine-glutamine reiterations (especially those between residues 61–102) constitute either statistically significant clusters of positively charged residues or statistically significant periodic patterns of positively charged, noncharged, noncharged residues as defined by Karlin et al. (60). These charge forms are notable because they are prominent among cellular proteins involved in regulatory functions, including nuclear transcription factors, steroid and thyroid hormone receptors, nuclear proto-oncogene products, developmental control proteins, heat shock proteins, and some transmembrane proteins. However, they are uncommon among most enzymes and constitutively expressed proteins. Transcription factors often carry clusters of positive net charge, and these clusters consisting of basic residues are often associated with DNA binding domains (23). The putative NLS (amino acids 17–20) is similar to that first identified in the large T-antigen of simian virus 40 (61, 62) versus the bipartite NLS first identified in the nonnucleic acid binding protein nucleoplasmin (63, 64). Residues 17–20 satisfy the requirements of a minimal NLS by having the tetrameric sequence K(R/K)X(R/K) in which X is K, R, P, V, or A but not N (47). The inclusion of amino acids 21 and 22 to the putative NLS allows it to be defined as a core NLS of the “highly basic” category (65). The flanking of the putative NLS with sites for possible serine phosphorylation suggests that the nuclear targeting of GRP-1 may be modulated by various stimuli or metabolic processes (66). The presence of four ATTTA consensus sequences in the 3′-untranslated region
Cloning and Expression of Glutamine Repeat Protein-1

25521

of GRP-1 cDNA suggests that the expression of GRP-1 is very highly controlled. Indeed, these AU-rich sequences are found in mRNAs of cytokines, oncogenes, transcription factors, and primary response genes (51, 67) and are implicated in mediating the degradation of mRNA (51, 68) and regulating the efficiency of translation (69, 70).

In vitro analyses proved that the cloned GRP-1 cDNA can be transcribed and translated, and they confirmed the open reading frame. However, SDS-PAGE analysis of the translation products detected several bands of protein with molecular weights greater than the calculated molecular mass of GRP-1 of approximately 21.6 kDa. In addition, similar bands were detected in Western blot analyses of extracts from ANA-1 macrophages or 3T3 fibroblasts transfected with GRP-1 cDNA. We suspect that the bands represented anomalous migration of GRP-1 in SDS-PAGE that might have been due to the very high positive charge (predicted pI 10.67) of GRP-1 or some structural feature imposed by its unusually high content of glutamines and histidines. The lower bands might have represented proteolytic cleavage products or various conformers of GRP-1. In related studies, GRP-1 was expressed in E. coli as a glutathione S-transferase fusion protein. The fusion protein had a molecular weight of approximately 48,000–50,000 according to SDS-PAGE. Cleavage of the fusion protein with thrombin generated glutathione S-transferase (26 kDa) and several bands of protein with molecular weight of approximately 28,000–35,000. Overall, these results verify the open reading frame in GRP-1 cDNA, demonstrate that the multiple bands detected by SDS-PAGE represented GRP-1, and confirm the specificity of the antipeptide serum.

Consistent with the identification of a NLS in the deduced amino acid sequence of GRP-1, GRP-1 was detected in the nucleus of ANA-1 macrophages. Although GRP-1 was localized exclusively in the nucleus of ANA-1 macrophages, it was found in both cytosolic and nuclear extracts of transfected 3T3 fibroblasts. The presence of GRP-1 in the cytosol of transfected fibroblasts most likely reflected its overexpression and subsequent saturation of the nucleus. However, it is also possible that migration of GRP-1 to the nucleus requires phosphorylation or dephosphorylation limited in the transfected cells compared with ANA-1 macrophages. Ongoing studies are addressing possible mechanisms that control nuclear localization of GRP-1.

Southern blot analyses of restriction enzyme-digested mouse genomic DNA showed that the full-length GRP-1 cDNA detects many bands of cross-hybridizing DNA. Similarly, Northern blot analyses of ANA-1 macrophage total RNA with full-length GRP-1 cDNA resulted in the detection of many bands of cross-hybridizing mRNA. These findings are consistent with the fact that many genes contain opa sequences. The use of a GRP-1-specific cDNA probe without opa sequences resulted in the detection of only one band of DNA in Southern blots and only two bands of mRNA in Northern blots. These results show that GRP-1 is encoded by a single-copy gene in the mouse and suggest that the two mRNAs derive from alternative splicing of a single primary RNA transcript.

Reverse transcription-PCR analyses were performed on total RNA from various cell types to begin to examine the distribution of GRP-1 mRNA. Constitutive levels of GRP-1 mRNA were detected in oncogene-immortalized macrophages and thioglycollate-elicited normal peritoneal exudate macrophages. However, in contrast to reverse transcription-PCR analysis, Northern blot analysis failed to demonstrate constitutive expression of GRP-1 mRNA in normal thioglycollate-elicited peritoneal macrophages. Because normal tissue macrophages do not proliferate and weakly express GRP-1 mRNA compared with immortalized ANA-1 macrophages, it is possible that augmented expression of GRP-1 is associated with enhanced cell cycle progression and proliferative capacity. The monocyte-macrophage cell line P388D1 and the myeloblast cell line M1 failed to express detectable levels of GRP-1 mRNA constitutively. It would be interesting to determine whether agents that cause M1 to become macrophage-like also induce expression of GRP-1 mRNA, and conversely, whether overexpression of GRP-1 in M1 results in macrophage differentiation. IFN-γ alone, LPS alone, and IL-2 alone augmented GRP-1 mRNA expression in ANA-1 macrophages. Because these agents have been shown to regulate many macrophage functions, including ANA-1 macrophage gene expression and tumoricidal activities (4, 45, 71), augmented expression of GRP-1 may mediate, at least in part, the effects of these agents or be associated with the activation of macrophages. Time-course studies showed that the effects of LPS or IFN-γ were not detectable after 1 h, were weakly detectable after 6 h, and were maximally detectable after 16 h. These results may reflect indirect effects of LPS and IFN-γ. Alternatively, they may reflect the ability of LPS and IFN-γ to stabilize GRP-1 mRNA. As mentioned above, there are several copies of ATTTA sequences in the 3′-untranslated region of GRP-1 cDNA that may render the GRP-1 gene susceptible to posttranscriptional regulation. Treatment of ANA-1 macrophages with the protein synthesis inhibitor cycloheximide augmented GRP-1 gene expression. Although this finding is consistent with the hypothesis that GRP-1 is posttranscriptionally regulated by LPS and IFN-γ, further studies are necessary to determine the mechanism(s) by which GRP-1 synthesis is regulated. Preliminary experiments have not shown increases in GRP-1 protein levels in ANA-1 macrophages after treatment with IFN-γ or LPS. We are examining the stability of GRP-1 and exploring further whether GRP-1 protein levels can be augmented by various agents.

The most compelling features of GRP-1 that suggest it is a transcription factor are its nuclear localization and its high percentage of glutamines (50%). It is well known that glutamine-rich regions constitute one of the three major classes of activation domain (the others are acidic regions and proline-rich regions). Glutamine-rich regions are required for the activity of transcription factor Sp1, and the overall glutamine content of these regions versus their primary sequence confers the ability of Sp1 to activate transcription (31). Indeed, polyglutamine alone can support transcriptional activation (35). Activation domains mediate interactions among various transcription factors that ultimately activate RNA polymerase II and gene transcription. Therefore, it is notable that Perutz et al. (26, 27) have proposed that glutamine repeats can function as polar zippers and cause homodimerization or heterodimerization of proteins. In their model, polypeptides with continuous stretches of glutamines form pleated antiparallel β-sheets joined by hydrogen bonds between their main chain amides and by hydrogen bonds between their side chains on either side of the sheet (27). Experiments confirmed that polyglutamine can cause protein oligomerization, and demonstrated that glutamine repeats form β-sheeted sheets (28).

To begin to study the function of GRP-1, we examined whether it can affect the activity of a transcription factor that uses glutamine-rich activation domains. Overexpression of GRP-1 decreased Sp1-driven CAT expression in transiently transfected 3T3 fibroblasts. In addition, overexpression of GRP-1 decreased the expression of SV40-luciferase, which is also regulated by Sp1 (72). There are several possible mechanisms by which repression might have occurred including the following: direct repression, competitive binding, sequestration
of unbound Sp1, and/or the process of quenching in which GRP-1 would bind to the activation domain of bound Sp1 and neutralize its activity. Although histidine-richness has been suggested to play a role in repression (73), GRP-1 does not contain regions high in alanine, proline, or glycine that are typical repressor domains (73–76). Therefore, the latter two mechanisms seem most plausible and might have resulted from glutamine-mediated interactions between GRP-1 and Sp1. Because the Sp1-CT construct only consists of Sp1 sites and a TATA element upstream of the CAT gene, the only other likely targets of GRP-1 were the basal transcription factors. In fact, TATA-binding protein and TAF110 are known targets of glutamine-rich transcriptional activators (77, 78). The inhibitory effects of GRP-1 also might have been due to squelching. Squelching is a phenomenon in which high concentrations of a transcriptional activator (or an activating region) can cause nonspecific repression via sequestration and titration of a component of the general transcription complex (79). It is not known whether GRP-1 can nonspecifically repress DNA via its motif, similar to GRP-1 in that it is glutamine (53%)- and histidine (25%)-rich and consists of repeating patterns of polyglutamine and interspersing histidines (81).

There are several mechanisms by which GRP-1 may activate gene transcription (82). It is not known whether GRP-1 can bind to DNA in a sequence-specific manner as either a monomer or a homodimer. Analysis of the deduced amino acid sequence of GRP-1 did not identify a typical DNA binding domain such as a helix-turn-helix, a zinc finger, or a basic DNA binding motif associated with either a leucine zipper or a helix-loophelix. However, GRP-1 may bind to DNA via an unknown DNA binding domain. Although most transcription factors bind to DNA independently, others (e.g. Fos and Myc) bind to DNA after dimerization with other factors (e.g. Jun and Max, respectively). Therefore, GRP-1 may bind to DNA as a complex with another known or unknown transcription factor. The assembly of such a DNA binding complex may occur via a polyglutamine-type polar zipper. Alternatively, GRP-1 may activate transcription by functioning as an adapter molecule or coactivator that bridges DNA-bound activators to the basal transcriptional complex. The accumulating evidence that polyglutamine and glutamine-rich regions can act as protein interaction motifs supports our hypothesis that GRP-1 is a transcription factor able to activate and/or repress gene expression.

Acknowledgments—We thank Drs. Howard Young and John Ortaldo for reviewing the manuscript. We also thank Dr. Luigi Varesio for his support and encouragement.

REFERENCES
1. Wiltrout, R. H., and Varesio, L. (1990) Immunopharmacology: The role of Cells and Cytoxins in Immunity and Inflammation (Opperhijn, J. J., and Shapira, M., eds) pp. 365–385, Oxford University Press, New York.
2. Blasi, E., Mathieson, B. J., Varesio, L., Cleveland, J. L., Borchert, P. A., and Rapp, U. (1985) Nature 318, 667–670
3. Blasi, E., Radzioch, D., Durum, S. K., and Varesio, L. (1987) Eur. J. Immunol. 17, 1431–1488
4. Cox, G. W., Mathieson, B. J., Gindino, L., Blasi, E., Radzioch, D., and Varesio, L. (1989) J. Natl. Cancer Inst. 81, 1492–1496
5. Kishimoto, A., Nishiyama, K., Nakanishi, H., Uratsuji, Y., Nomura, H., Takeyama, Y., and Nishizuka, Y. (1985) J. Biol. Chem. 260, 12492–12499
6. Pinna, L. A. (1990) Biochem. Biophys. Acta 1054, 267–284
7. Shaw, G., and Kamen, R. (1986) Cell 46, 659–667
8. Pinna, L. A., and Varesio, L. (1992) Glutamine Repeat Protein-1: Cloning and Expression of Glutamine Repeat Protein-1. Academic Press, Inc., San Diego.
Cloning and Expression of Glutamine Repeat Protein-1

69. Krupys, V., Wathelet, M., Poupart, P., Contreras, R., Fiers, W., Content, J., and Huez, G. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6030–6034
70. Han, J., Brown, T., and Beutler, B. (1990) J. Exp. Med. 171, 465–475
71. Cox, G. W., Mathieson, B. J., Giardina, S. L., and Varesio, L. (1990) J. Immunol. 145, 1719–1726
72. Gidoni, D., Kadonaga, J. T., Barrera-Saldana, H., Takahashi, K., Chambon, P., and Tjian, R. (1985) Science 230, 511–517
73. Han, K., and Manley, J. L. (1993) Genes Dev. 7, 491–503
74. Catron, K. M., Zhang, H., Marshall, S. C., Inostroza, J. A., Wilson, J. M., and Abate, C. (1995) Mol. Cell. Biol. 15, 861–871
75. Licht, J. D., Grossel, M. J., Figge, J., and Hansen, U. M. (1990) Nature 346, 76–79
76. Han, K., and Manley, J. L. (1993) EMBO J. 12, 2723–2733
77. Gill, G., Pascal, E., Tseng, Z. H., and Tjian, R. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 192-196
78. Emmi, A., Greenblatt, J., and Ingles, C. J. (1994) Mol. Cell. Biol. 14, 1582–1593
79. Gill, G., and Ptashne, M. (1988) Nature 334, 721–724
80. Dubin, R. A., and Oster, H. (1994) Mol. Endocrinol. 8, 1182–1192
81. Whitfield, L. S., Lovell-Badge, R., and Goodfellow, P. N. (1993) Nature 364, 713–715
82. Latchman, D. S. (1995) Eukaryotic Transcription Factors, pp. 239–278, Academic Press, Inc., San Diego