Molecular Analysis of a Major Antigenic Region of the 240-kD Protein of Mi-2 Autoantigen

Qun Ge,‡ David S. Nilasena, Charles A. O'Brien, Mark Barton Frank,‡ and Ira N. Targoff‡*

**Department of Medicine, University of Oklahoma Health Sciences Center and Veterans Affairs Medical Center; and the ‡Arthritis/Immunology Section, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104*

Abstract

Anti-Mi-2 autoantibody is strongly associated with dermatomyositis and found in sera of 20% of patients. Mi-2 antigen contains at least eight components and previous evidence suggested that the 240-kD protein was the antigenic component for at least some sera. In this study, anti-Mi-2 patient sera were used to screen human thymocyte and HeLa cell λgt11 expression libraries, and two clones from each had plaques specifically reactive with anti-Mi-2 sera. Studies with affinity-purified antibody supported the identification of the clones. All of 44 anti-Mi-2 sera reacted with the plaques, but none of 44 control sera reacted significantly. The cDNAs were identical, and full sequencing of one revealed an open reading frame spanning a 1,054-bp insert. Rescreening the library with the cDNA yielded a 1,589-bp cDNA that continued the open reading frame. The Mi-2 cDNA hybridized to a single 7.5–8.0 kb mRNA of HeLa cells, by Northern blot. Rabbit antiserum directed at a portion of the cDNA product reacted with HeLa 240-kD Mi-2 protein. The sequence was notable for four potential zinc-fingers and several charged regions. The protein encoded by the cDNA produced in vitro reacted with only one of five of the Mi-2 sera. These findings indicate that the Mi-2 240 kD is a novel protein that is antigenic for all Mi-2 sera, and strongly suggests that a major common epitope is conformational in nature. (*J. Clin. Invest.* 1995. 96:1730–1737.) Key words: myositis • dermatomyositis • autoantibodies • nuclear proteins • autoantigens

Introduction

The idiopathic inflammatory myopathies polymyositis and dermatomyositis (DM)[1] are associated with serum autoantibodies to intracellular proteins in 60–80% of patients (1). Certain of these autoantibodies are found exclusively or predominantly in myositis patients, referred to as myositis-specific autoantibodies. The disease specificity of these autoantibodies makes them of special interest, since it suggests a fundamental relationship to etiologic and pathogenetic factors in the disease (2). The most common of these overall is anti-Jo-1 (3), which reacts with histidyl-tRNA (transfer RNA) synthetase (4). Several other myositis-specific autoantibodies are directed at other aminoacyl-tRNA synthetases (5–7) or other cytoplasmic antigens such as the signal recognition particle (8).

One myositis-specific autoantibody, however, anti-Mi-2, reacts with a nuclear antigen that is absent from the cytoplasm and the nucleoli (9). Anti-Mi-2 is also unique in that it is almost exclusively associated with DM rather than polymyositis (PM) (9, 10), and has been found in 15–20% of DM patients (9). This is the only autoantibody for which such a close association with DM has been established. Patients with anti-Mi-2 often have florid cutaneous manifestations including typical Gottron's papules and heliotrope lesions, and involvement of the base of the neck and upper back ("V" and "shawl" patterns) (10). Anti-Mi-2 is found in all classes of DM patients including children (9, 11, 12), but is not associated with the constellation of findings seen with antisynthetases (10), such as interstitial lung disease, arthritis, or Raynaud's phenomenon.

Despite these interesting features of the autoantibody, little is known about the antigen to which it is directed. By immunoprecipitation (IPP) from [^{35}S]methionine-labeled HeLa cells, all anti-Mi-2 sera show a series of at least eight polypeptides that include a prominent band of ~ 240 kD, along with weaker bands of 200, 150, 75, 65, 63, 50, and 34 kD (11), raising the possibility of a macromolecular complex. The relative prominence of 240-kD protein suggests that it is directly reactive, but reaction of anti-Mi-2 autoantibodies by immunoblot has been difficult to demonstrate. Recently, however, about half of anti-Mi-2 sera were shown to react by blot with the HeLa Mi-2 240-kD protein, and a smaller proportion with other Mi-2 proteins, by using an Mi-2-enriched antigen prepared by IPP (13). It was not known whether the other sera also recognize the 240-kD protein under other conditions. No nucleic acid has been associated with Mi-2 antigen, and there is no information regarding its function.

In this study, the 240-kD protein was confirmed to be the major Mi-2 antigenic protein, and a major immunoreactive region was identified and characterized by isolation and characterization of cDNA clones encoding a portion of the protein.

Methods

Sera. Serum samples with anti-Mi-2 and control sera without anti-Mi-2 were obtained from the serum bank at the Oklahoma Medical Research Foundation, and included some anti-Mi-2 sera that were re-

David S. Nilasena is currently a medical consultant with Health Care Financing Administration, Division of Health Standards and Quality, Medical Review Branch, Region VI, Dallas, TX 75217. Charles A. O'Brien's current address is Division of Endocrinology and Metabolism, University of Arkansas for Medical Sciences, Little Rock, AK 72205.

Address correspondence to Ira N. Targoff, Arthritis/Immunology Section, Oklahoma Medical Research Foundation, 825 NE13th, Oklahoma City, OK 73104. Phone: 405-271-7395; FAX: 405-271-4110.

Received for publication 7 April 1994 and accepted in revised form 8 June 1995.

1. *Abbreviations used in this paper:* aa, amino acids; DM, dermatomyositis; IPP, immunoprecipitation; IPPs, immunoprecipitates; tRNA, transfer RNA.

J. Clin. Invest.

0021-9738/95/10/1730/08 $2.00
Volume 96, October 1995, 1730–1737

ferred from other centers as part of other studies. 44 sera were identified as anti-Mi-2 positive by double immunodiffusion with a standard serum identical with the original prototype (9), and confirmed using IPP from [^{35}S]methionine-labeled HeLa cells (13). 42 of 44 were from DM patients (1 amyopathic), and 2 from PM patients. In a previous study (13), 22 of the anti-Mi-2 sera used here were determined to be reactive by immunoblot against HeLa 240-kD protein, and 22 were negative. Control sera included 22 sera from normal laboratory workers, and 22 sera (19 from myositis patients) with other autoantibodies, including anti-Jo-1, anti-PL-7, anti-PL-12, anti-OJ, anti-PM-Scl, anti-Ro, anti-La, anti-RNP, and anti-Sm.

Isolation of clones. Two λgt11 cDNA libraries, one derived from human thymocytes and the other from HeLa cells (Clontech, Palo Alto, CA), were independently screened with anti-Mi-2 patient plasma from patient ML, from the method of Young and Davis (14). Sera were either used directly at 1:500 dilution in Tris-buffered saline with 0.05% Tween 20, or were used at 1:100 dilution after preincubation with *Escherichia coli* proteins and plaques of wild-type λgt11 phage on nitrocellulose filters to remove antibodies to *E. coli* proteins. 5% nonfat milk in serum dilution buffer was used as blocking agent. Goat anti-human IgG/alkaline phosphatase (Sigma Chemical Co., St. Louis MO), was used as conjugate, and bromo-chloro-indolyl phosphate/nitroblue tetrazolium as substrate. Reactive clones were plaque purified and further tested as below.

Clone L4 cDNA was labeled with ^{32}P and used as a probe to screen a human thymocyte "5′ stretch" λgt11 cDNA library (Clontech). After hybridization with transferred plaques, filters were developed by autoradiography, and positive clones were plaque purified. Those with the longest cDNA inserts were selected and further characterized.

Testing of clones. Further serologic testing of clone L4 was performed using a plaque-test in which *E. coli* Y1090 was incubated with a 1:1 mixture of purified reactive phage and wild-type phage, so that anti-Mi-2 sera would react with 50% of plaques, with 50% as negative controls (15). Portions of nitrocellulose were incubated separately with each test serum and developed as described above for screening. Affinity purification of antibodies reactive with plaque proteins of L1 or L4 recombinant λgt11 was accomplished as previously described (15). Essentially, 100 μl anti-Mi-2 plasma was diluted 1:100 and absorbed extensively with *E. coli* lysate, and incubated with plaques from Mi-2 recombinant or control wild-type clones. Bound antibody was eluted with 0.2 M glycine-HCl pH 2.6 and immediately neutralized. 500 μl of eluate was obtained and used undiluted for subsequent experiments. Fusion proteins were further analyzed by immunoblot of lysates of five 150-cm^2 *E. coli* Y1090 plates with the L4 Mi-2 clone (plate lysate) using anti-Mi-2 sera, control sera, and a mouse anti-β-galactosidase monoclonal antibody (anti-β-gal; Boehringer Mannheim Corp., Indianapolis, IN), and by immunoprecipitation from plate lysate using these sera. The immunoprecipitates were analyzed by immunoblot developed with the monoclonal to detect the fusion protein (15).

cDNA characterization. The EcoRI insert from each isolated clone was electroeluted from a 1% agarose gel and subcloned into pUC19 (clones L1-4) or pUC18 (clone C4) vector (16). The nucleotide sequence was determined by the dideoxynucleotide chain-termination method (17) using modified T7 DNA polymerase (Sequenase II) as per the manufacturer's recommendation (U.S. Biochemical Corp., Cleveland OH). Initial double-stranded sequencing of clones L1, L3, and L4 was performed using mini-prep plasmid DNA (18). For further sequencing, clone L4 cDNA was digested with PstI, BanII, StyI, and TaqII, and resulting overlapping fragments were subcloned into M13mp18 vector. Whole C4 cDNA was subcloned into M13mp18, and sequenced using synthesized gene-specific primers. cDNA sequences were analyzed using the Sequence Analysis Software Package (Genetics Computer Group, Inc., Madison, WI) run on a VAX 4600 computer. Database searches were performed at the National Center for Biotechnology Information using the BLAST network service (19). Regions of charged runs and clusters, repetitive structures, and periodic motifs in proteins were analyzed using the Statistical Analysis of Protein Sequences program, generously made available to us by Dr. Volker Brendel at Stanford University (Standford, CA) (20).

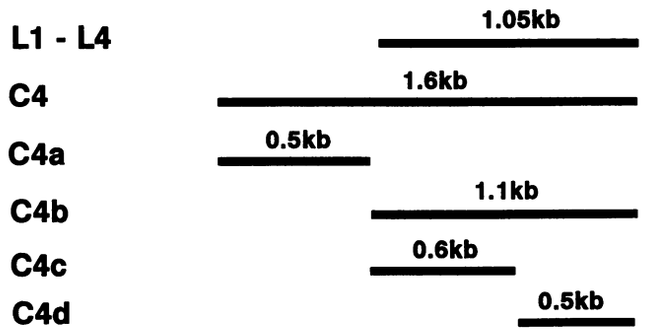

Figure 1. Diagram of Mi-2 cDNAs and fragments. Clones L1 to L4 were isolated by immunoscreening; L1 and L2 from the thymocyte library, L3 and L4 from the HeLa library. C4 was isolated by screening with L clones. Fragments C4a and C4b were derived by digestion of whole C4 with BamHI and HindIII. Fragments C4c and C4d were derived by PCR.

Northern blots. HeLa cell RNA was prepared for Northern blot as described previously (15), including total RNA (21) and poly-(A)+ RNA (MicroFastTrack; Invitrogen, San Diego, CA). RNA was electrophoresed in a denaturing 1% agarose gel with 2 M formaldehyde and transferred to nylon membrane (Nytran; Schleicher & Schuell, Inc., Keene, NH). Size markers were synthetic poly (A)-tailed RNAs from 0.24 to 9.5 kb (RNA ladder; Gibco BRL, Gaithersburg, MD).

Hybridization. Radiolabeled cDNA probes were prepared by random priming using a commercial kit (U.S. Biochemical Corp). The filters were prehybridized at 42°C for 4 h in a 2× SSC solution, then hybridized with labeled probe for 16 h at 42°C. The filters were then washed three times with 2× SSC for 5 min and once with 0.2× SSC for 15 min at 65°C with two changes before autoradiography (16, 22).

Expression of C4 cDNA fragments. The cDNA of clone C4 was amplified by PCR using gene-specific oligonucleotide primers (5′-TCCAGCAGAAGCAAAAGCAAGTG-3′ and 5′-TTATCCACACTG-TGGTTGATG-3′), with BamHI and HindII sites added to facilitate subcloning into pQE expression vector. This PCR product was digested with BamHI and HindIII to yield a 0.5-kb 5′ fragment (C4a; Fig. 1) and 1.1-kb 3′ portion (C4b). Two fragments that covered the 1.1-kb 3′ portion (C4b) were separately prepared by PCR with gene-specific primers with BamHI or HindIII sites attached, including a 0.6-kb central C4 fragment (C4c; Fig. 1) (5′-TGATGGACGCAAGAAGC-3′ and 5′-TTAGACAATGAATGTGG-3′) and a 0.5-kb 3′ fragment (C4d; Fig. 1)(5′-ACATTCATTGTCTAAAC-3′ and 5′-ATCCACACTGTGGTT-GATG-3′). The individual isolated DNA fragments were digested with BamHI and/or HindIII, subcloned into pQE plasmid, and transformed into *E. coli* M15/pREP4 using the Qiaexpressionist system (Qiagen Inc., Chatsworth, CA). The protein product was solubilized in guanidine/HCl and purified using Ni-NTA affinity chromatography with a urea gradient according to manufacturer's instructions. Urea was reduced to a concentration of 0.1 M by dialysis.

Attempts to produce a λgt11 lysogen in Y1089 were performed as described previously (15). Expression of the intact C4 cDNA was attempted as described above using the Qiaexpressionist system, as well as the ThioFusion Expression System and Baculovirus Protein Expression System (Invitrogen) according to manufacturer's instructions.

Production of rabbit antiserum. A New Zealand white male rabbit was immunized with 500 μg of the purified 20-kD expressed product ("NH_2-terminal protein") of the 5′ fragment of Mi-2 clone C4 (C4a; Fig. 1), in 0.5 ml phosphate buffered saline and an equal volume of complete Freund's adjuvant, then boosted with 500 μg antigen in incomplete Freund's adjuvant on the second and fourth weeks. Rabbit serum activity was tested by immunoblot against NH_2-terminal protein.

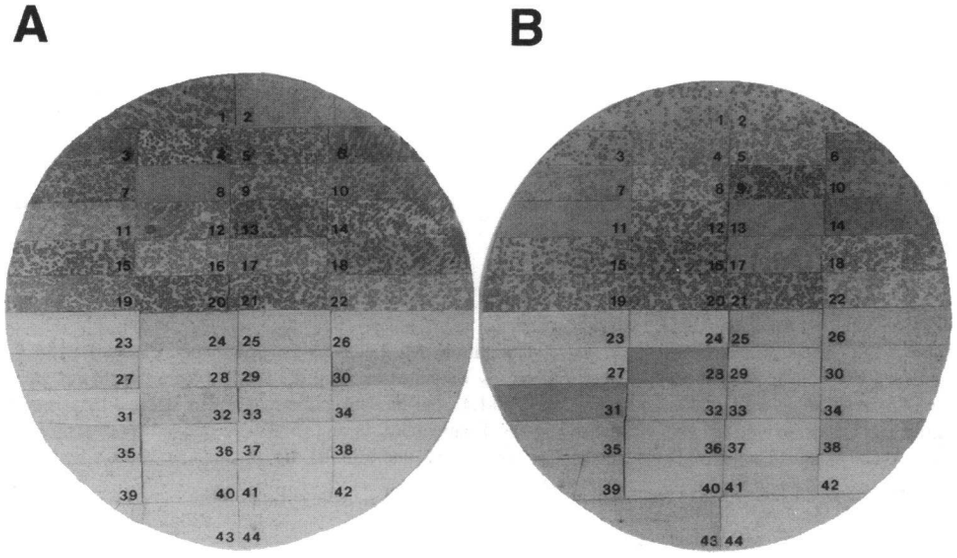

Figure 2. Immunoreactivity of anti–Mi-2 sera with plaques of clone L4. Nitrocellulose filters were used to blot a 1:1 mixture of IPTG-induced plaques from recombinant L4 and wild-type λgt11 bacteriophage. Each numbered nitrocellulose section was developed using a different serum, preabsorbed with lysed *E. coli,* and diluted 1:500, as follows: Sections A1–A22 = anti–Mi-2 sera that had reacted by immunoblot with 240-kD protein (13); B1–B22 = anti–Mi-2 sera that did not immunoblot 240 kD; A23–A33 and B23–B33 = normal sera; A34–A44 and B34–B44 = anti–Mi-2-negative sera with other autoantibodies (see Methods).

Immunoblotting. Immunoblotting was performed using as antigen either IPPs from HeLa extract (or *E. coli* lysate) prepared with anti–Mi-2 patient serum, NH_2-terminal protein (purified or unpurified), or in vitro translation lysates. To prepare the IPPs, 20 μl protein A-Sepharose was coated with 15 μl serum, and incubated with extract from ½ of a 150 cm^2 flask of HeLa cells. After 10% SDS-PAGE, proteins were transferred to nitrocellulose. Blot developing buffer, blocking solution, conjugate, and substrate were as described above for antibody screening of libraries. When IPPs were used as antigen, protein A-Sepharose was cross-linked with dimethyl-suberimidate before incubation with HeLa extract, using the Affinica Antibody Orientation Kit (Schleicher & Schuell, Inc.).

In vitro transcription/translation. Clone C4 cDNA was prepared by PCR with introduction of start and stop codons, and subcloned into the BamHI site of pGEM-3Zf plasmid (Promega Corp. Madison, WI). The correct reading frame was confirmed by nucleotide sequencing through the cloning site. In vitro translation was accomplished using the TnT SP6 Coupled Rabbit Reticulocyte Lysate System (Promega Corp.) with a plasmid containing C4 insert as transcription template. Protein translation product with labeled or unlabeled [^{35}S]methionine was detected by autoradiography or immunoblot.

Other methods. IPP from [^{35}S]methionine-labeled HeLa cells, was performed as described previously (7, 23). Each immunoprecipitate, used for one lane, was prepared using 20 μl protein A-Sepharose coated with 15 μl patient serum, and incubated with extract from 1/10 of one 25 cm^2 flask of HeLa cells. Indirect immunofluorescence was performed using commercial HEp-2 slides (Protrac, Kerrville, TX).

Results

Identification of cDNA clones. From initial screening of ~ 1,000,000 plaques, two clones were identified from the HeLa library (L1–L2) and two from the thymocyte library (L3–L4) that reacted strongly with 6 anti–Mi-2 sera. Clones L1 and L2 each reacted weakly with only one of eight control sera, and clones L3 and L4 were negative with all eight. Clone L4 was further tested for reaction with anti–Mi-2 sera from a total of 44 different patients (Fig. 2), including 22 that had reacted by blot with the 240-kD protein, and 22 that had not (13), 44 anti–Mi-2 negative control sera were also tested, including 22 from normals and 22 from patients with myositis or with other autoantibodies. All 44 anti–Mi-2 sera were positive, and all 44 control sera were negative (Fig. 2).

Affinity-purified antibodies from plaques of clone L4 strongly reacted with plaques of each of the four clones, and showed similar reaction to that of whole anti–Mi-2 serum by indirect immunofluorescence on HEp-2 cells. Eluates from wild-type plaques were negative in both assays. By IPP from HeLa cell extract, affinity-purified antibodies from clones L1 and L4, but not wild-type eluates, immunoprecipitated the 240-

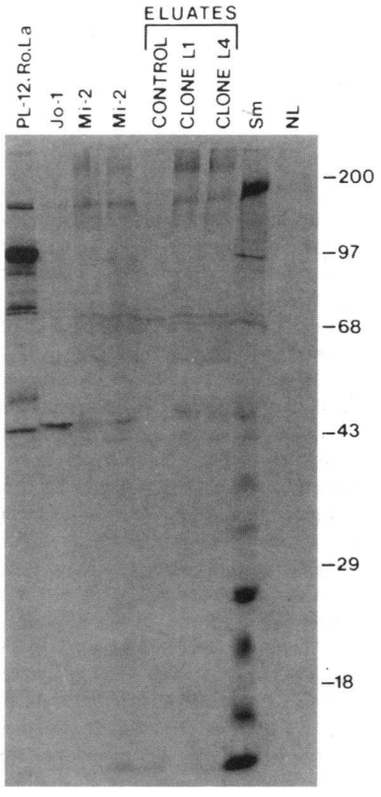

Figure 3. Immunoprecipitation with affinity-purified antibodies. 10% SDS-PAGE of HeLa immunoprecipitates prepared with patient sera with autoantibodies to the antigens indicated (*PL-12, Ro, La, Jo-1, Mi-2, or Sm*), normal serum (*NL*), or affinity-purified antibodies (*ELUATES*) from plaques of clone L1, clone L4, or wild-type λgt11 (*CONTROL*). For sera, 15 μl whole serum was used to coat the protein A-Sepharose; for eluates, 200 μl was used undiluted, obtained as in Methods. The 240-kD and the 65- and 63-kD Mi-2 components are visible in the anti–Mi-2 lanes and the L clone lanes. Positions of molecular weight markers in kD are at right.

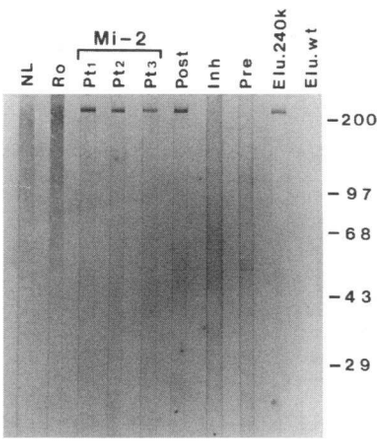

Figure 4. Immunoblot of HeLa anti-Mi-2 immunoprecipitates. Immunoprecipitates from HeLa cell lysate were prepared using anti-Mi-2 patient serum, electrophoresed in 8% SDS-PAGE, and transferred to nitrocellulose. The results of one immunoprecipitation, prepared by coating 20 μl protein A-Sepharose with 15 μl serum and incubating with extract of 1/2 of a 150 cm^2 flask of HeLa cells, was applied to each lane. Lanes were developed as for immunoblot with: anti-Mi-2 patient sera (*Pt1, Pt2, Pt3*); anti-Ro control sera; normal serum (*NL*); affinity-purified antibodies prepared by incubation of anti-Mi-2 sera with plaques of clone L4 (*Elu.240K*) or wild-type λgt11 (*Elu.wt*), followed by elution; preimmune rabbit serum (*Pre*), postimmunization antiserum (*Post*), or antiserum preincubated with NH_2-terminal protein (*Inh*). Patient sera were used at 1:200 dilution, rabbit sera at 1:2,000 dilution, and affinity purified antibodies were 200-μl eluates used undiluted as in Methods. Positions of molecular weight markers are at right.

kD protein seen with whole anti-Mi-2, and faintly the 65- and 63-kD proteins (Fig. 3). Affinity-purified antibodies also reacted by immunoblot with the 240-kD protein but not other Mi-2 proteins when tested against anti-Mi-2 IPPs (Fig. 4), while eluates from wild-type did not react.

Despite strong reaction with plaques of clone L4, anti-Mi-2 sera did not show reaction with a fusion protein in plate lysates of *E. coli* Y1090 with L4 λgt11 by immunoblot. By IPP (not shown), anti-β-gal monoclonal showed a strong 111-kD protein with wild-type λgt11 plate lysate, along with a 94-kD apparent degradation product. With clone L4 λgt11 plate lysate, IPP with anti-β-gal showed disappearance of the intact β-gal, a substantially increased amount of the 94-kD protein, and a very faint new band at 114 kD, but not a band consistent with the expected whole fusion protein (156 kD), suggesting increased susceptibility of the fusion protein to degradation. No band consistent with an Mi-2 portion of the fusion protein was seen by protein stain or anti-Mi-2 blotting, suggesting further degradation of this portion.

cDNA characterization and sequencing. Clones L1-L4 each yielded similarly-sized inserts of 1.05 kb. More than 100 bp of sequence from both ends of clones L1, L3, and L4 were identical between clones. The remaining sequence was determined for clone L4 (Fig. 5 *A*), showing that the cDNA contained 1054 bp encoding 351 amino acids (aa) without a stop codon, a probable initiation site (24), or a poly-A tail, indicating that it was a portion of a larger cDNA. This was verified by Northern blot of labeled L4 cDNA against total HeLa cell RNA, which revealed a single band of ~ 7.5-8.0 kb (Fig. 6).

12 cDNA clones identified by probing a λgt11 thymocyte library with labeled L4 cDNA were then studied, and all 12 had inserts between 0.5 and 1.6 kb. Sequencing of both strands of the longest cDNA, clone C4, showed 1,589 bp (excluding 5′ and 3′ EcoRI-linkers), including the above L4 sequence and extending 525 bp beyond its 5′ end and 29 bp beyond its 3′ end (Fig. 5 *A*). C4 cDNA extended the L4 cDNA open reading frame, still without stop codon or probable start site, encoding 529 aa with predicted molecular mass of 59.3 kD.

Expression of cDNA fragments. An attempt was made to produce the L4 fusion protein by using a lysogen of L4 phage in *E. coli* Y1089, but surprisingly, > 99% of infected *E. coli* died, with surviving colonies showing altered morphology with ragged and pointed edges. Similarly, failure of attempted expression of C4, L4, or C4c cDNAs in several systems, with resulting deletions or other instability of the plasmid vector, indicated that the cDNA products were toxic to *E. coli.*

Successful expression of the 5′ and 3′ portions (C4a and C4d), but not the central portion, of the C4 cDNA was achieved, however. Anti-Mi-2 serum from patient VA, which had reacted strongly with the 240-kD Mi-2 protein by immunoblot of IPPs, reacted strongly with the 20-kD NH_2-terminal protein (Fig. 7), which was not present in the fusion proteins of L1-L4 λgt11 clones, although 11 other anti-Mi-2 sera tested did not react. Only 1 (not VA) of these 12 patient sera tested reacted with the COOH-terminal protein (product of the 3′ C4 cDNA fragment).

Rabbit antiserum. Rabbit antiserum raised against the 20-kD NH_2-terminal protein (product of C4a cDNA), but not preimmune rabbit serum, was strongly reactive with that protein (titer 1: > 100,000) both in *E. coli* lysate and after purification (Fig. 7). This serum, but not preimmune serum, reacted with the Mi-2 240-kD protein by immunoblot against anti-Mi-2 IPPs from HeLa cells, but did not react with other Mi-2 proteins (Fig. 4). The reaction with Mi-2 240 kD was completely blocked by preincubation of rabbit antiserum with purified C4 NH_2-terminal protein (Fig. 4, *Inh*), but, as expected, not with a control recombinant protein (PM-Scl 100 kD, [15]), providing further confirmation that the C4 NH_2-terminal protein represented a portion of Mi-2 240 kD.

Unlike patient sera, the rabbit antiserum showed very little activity by IPP, either being negative or, in some experiments, weakly positive for 240 kD, always negative for 150 kD, and equivocal for other Mi-2 components. This suggested that the portion of 240 kD protein corresponding to the C4 NH_2-terminal region may not be exposed in the native conformation. The rabbit serum was also negative by immunodiffusion against calf thymus extract. However, indirect immunofluorescence with rabbit antiserum showed strong nuclear staining, in a pattern similar to that of anti-Mi-2 sera, except for nonspecific cytoplasmic staining present also with preimmune serum.

Testing recombinant protein. After in vitro transcription and translation of the full C4 cDNA in a rabbit reticulocyte system, the postimmune rabbit serum and anti-Mi-2 patient serum VA reacted with a 60-kD protein that was identical in size to the ^{35}S-labeled translation product, of expected size for the C4 product, and present only in lysates translating the C4 clone, rather than control cDNA (Fig. 8). The other four anti-Mi-2 sera shown in Fig. 8, two of which were Mi-2 240 kD blot positive, did not react with the translation product by immunoblot. Control sera were also negative. Five additional anti-Mi-2 sera (two of which were 240 kD blot positive) were similarly tested in a separate experiment (not shown), and all were negative, while serum VA and postimmune rabbit serum were again positive.

Amino acid sequence analysis. The translated protein sequence of the C4 clone is indicated in Fig. 5 *A,* representing, as noted, only a portion of the coding region of the Mi-2 240-kD protein. Significant features included: *(a)* two regions rich

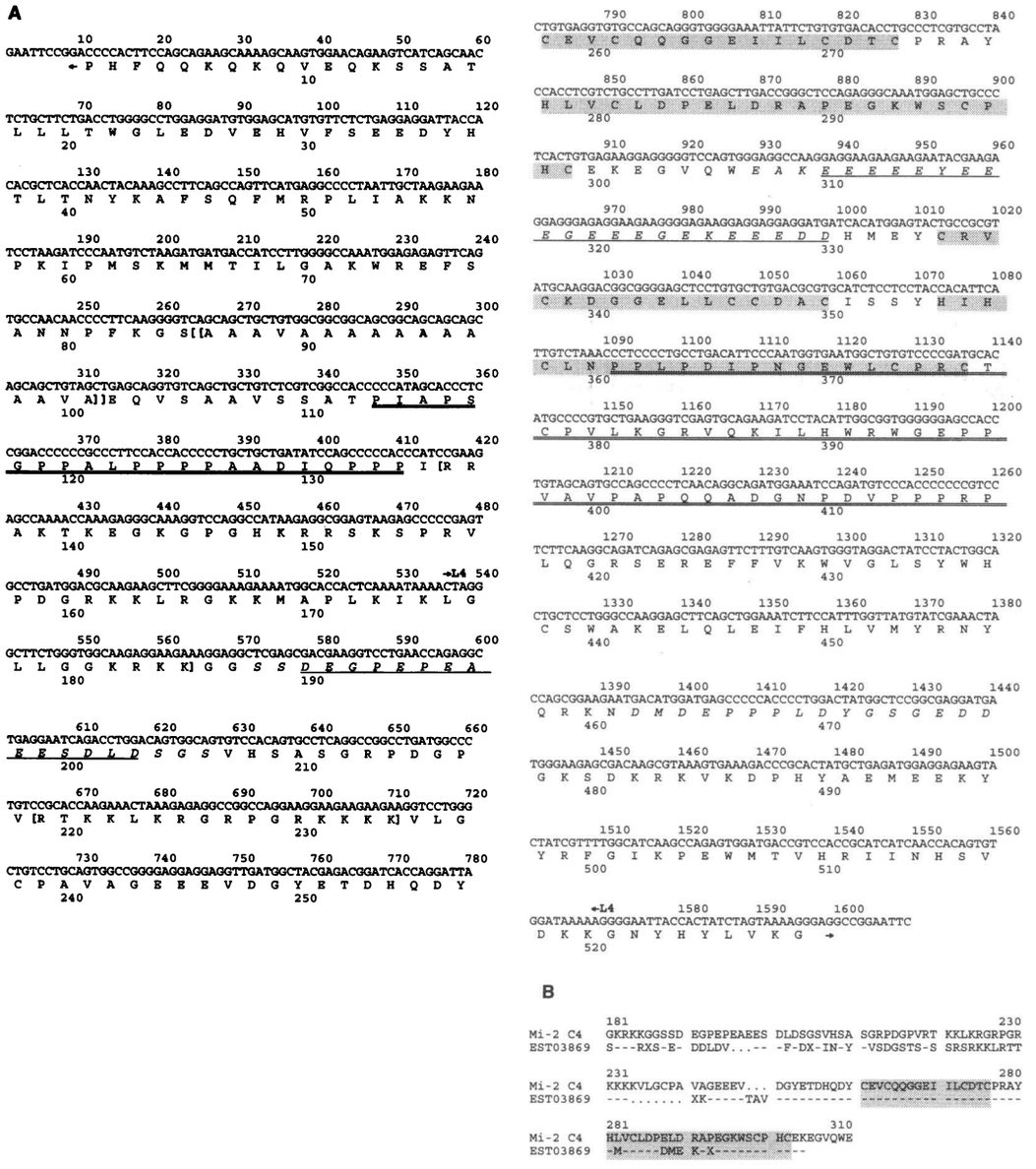

Figure 5. (A) Complete nucleotide and deduced amino acid sequence of C4 cDNA clone. Sequence includes 1609 total nucleotides (1589 excluding EcoRI linkers) encoding a predicted protein of 59.3 kD containing 529 amino acids. Underlined = acidic regions (aa 190–203, 310–330). Bracketed = basic regions (aa 136–185, 219–234). Double Bracket = alanine-rich region (aa 86–101). Double underline = proline-rich regions (aa 113–134, 361–417). Shaded = Potential zinc-finger regions (aa 258–273, 278–299, 335–350, 355–376). Italics = PEST/D regions (aa 188–206, 307–330, 462–477). ← and → = 5′ and 3′ EcoRI linkers of clone C4. →L4 and ←L4 = 5′ start and 3′ end sites of the identical sequence from clone L4 (nucleotides 536 and 1571). *(B)* Comparison of the translation of reported sequence of EST03869 (cDNA clone HFBDK47, accession #T05980, [25]) to the homologous region of the Mi-2 clone C4. Numbers are from the predicted amino acid sequence of Mi-2 240-kD protein from clone C4. Dashes indicate identity between the predicted EST and Mi-2 amino acid sequences. Dots represent gaps to optimize alignment. Shaded areas are potential Zn-finger regions.

in proline residues, one 11 of 22 (50%), the other 15 of 57 (26%); *(b)* one region rich in alanine residues, 14 of 16 (88%); *(c)* two regions rich in basic amino acids, one 22 of 50 (44% Arg+Lys), the other 11 of 16 (69% Arg+Lys;); *(d)* two regions rich in acidic residues, 8 of 14 (57% Asp+Glu) and 17 of 21 (81% Asp+Glu), *(e)* three regions with a composition consistent with PEST/D regions; and *(f)* four regions with a potential zinc-finger structure, arranged as two sets of two regions (Fig. 5 *A*). The two potential zinc fingers of each set were separated by four aa, and the sets were separated by a region that included the 21 aa acidic region, rich in glutamic acid. The two sets had 55% similarity to each other in sequence and spacing, as determined by dot plot analysis. The first potential finger of each set began and ended with two cystines separated by two aa (C-X-X-C) with an eight aa loop between, while the second potential finger of each set began with histidine and cystine separated by two aa (H-X-X-C), and ended with C-X-X-C, with a 14 aa loop.

The nucleotide and amino acid sequence of C4 cDNA were unique in the GenBank, EMBL and PIR databases. One ex-

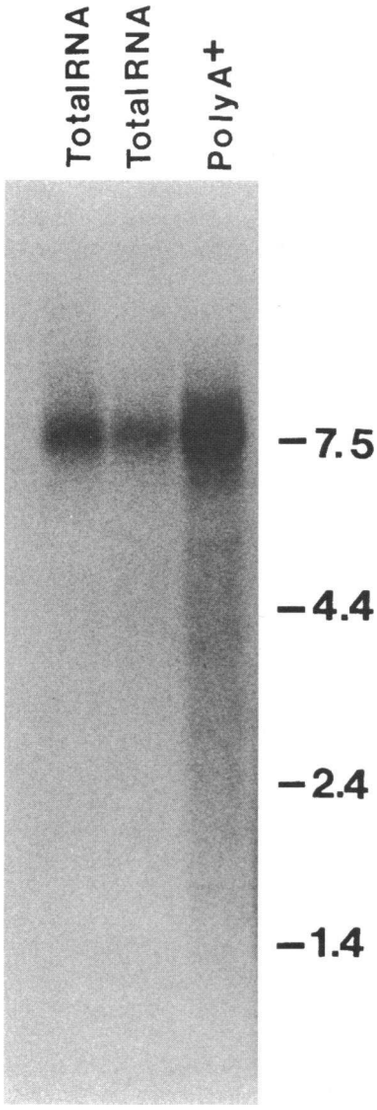

Figure 6. Northern blot. Total HeLa RNA of two different amounts (30 μg in left lane, 15 μg in center lane), and 8 μg of poly(A)+ HeLa RNA, were electrophoresed in denaturing 1% agarose gel with 2 M formaldehyde, transferred to Nytran, and hybridized with labeled L4 insert. The positions of synthetic poly(A)+ RNA size markers (see Methods) are shown at right.

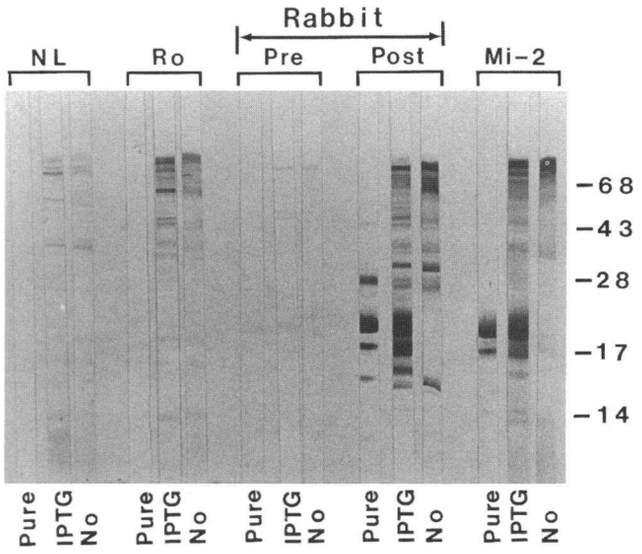

Figure 7. Immunoblot of recombinant NH_2-terminal protein. Each group of three nitrocellulose strips was developed with the serum indicated at top. The antigens included: *(No) E. coli* lysate with plasmid with no insert, induced by IPTG; *(IPTG)* lysate of induced *E. coli* with plasmid with the 5′ portion of C4 cDNA; *(Pure)* purified NH_2-terminal protein (20 kD). Sera, all used at 1:200 dilution, include: *(NL)* normal human serum; *(Ro)* anti-Ro serum; *(Mi-2)* anti-Mi-2 serum VA; *(Pre)* preimmune rabbit serum; *(Post)* postimmunization with NH_2-terminal protein. All sera recognize *E. coli* proteins, but the recombinant 20-kD protein, present only in induced lysate or purified preparation, is recognized only by serum VA and postimmune rabbit serum. The purified NH_2-terminal protein had a 28-kD contaminant recognized by the rabbit immunized with this preparation but not by anti-Mi-2 serum; the 17-kD protein, recognized by both patient and rabbit serum, may be a degradation product.

pressed sequence tag from a cDNA derived from human brain (EST03869, accession T05980, clone HFBDK47) (25) had 88% identity with Mi-2 in the Zn-finger region, and 44% identity and 94% similarity (with gaps for alignment) in the adjacent NH_2-terminal region. Two other proteins, human hrx and laminin *a* chain precursor, had significant identity (42 and 30%) and similarity (79 and 57%) that was limited to the Zn-finger regions.

Discussion

We have isolated a cDNA clone encoding a portion of the Mi-2 240-kD protein that was recognized by all tested patients with anti-Mi-2 autoantibodies. Three independent lines of evidence were used to establish the identity of the protein partially encoded by this cDNA as Mi-2 antigen. The first was serological testing, which demonstrated complete concordance of the presence or absence of anti-Mi-2 with reaction with the recombinant plaques. We have now tested anti-Mi-2 sera from over 60 patients, and there is still a 100% correlation. Second, affinity-purified antibodies from plaques of clone L4, but not wild-type, showed the same reactivity as whole anti-Mi-2 sera by indirect immunofluorescence and IPP, and bound the HeLa 240-kD protein by immunoblot of anti-Mi-2 IPPs, but not other Mi-2 components. Third, rabbit antiserum to the recombinant protein fragment reacted with HeLa 240-kD Mi-2 protein. Since the immunizing protein was derived from the 5′ extension present on C4 cDNA but not L4 cDNA, this experiment further demonstrated that the extended portion was also represented on the Mi-2 protein.

Autoantibody reactivity

Reaction with 240 kD. Although as noted, the 240-kD component of Mi-2 antigen had previously been suspected to be the major antigen (13), only 50% of sera had been shown to react with it. This study provided much stronger evidence, since 100% of the 44 anti-Mi-2 sera reacted by plaque test with the encoded portion of the protein. Failure to detect immunoblot reaction with the 240-kD protein by 50% of anti-Mi-2 sera in that study could have been due to exclusive reaction of those sera with conformational epitopes, inadequate amount of protein presented, or other factors. The finding that reaction with the region studied here appeared to be conformational is consistent with the former explanation. The uniform reactivity with the 240-kD protein does not exclude the possibility of additional reactivity of other Mi-2 components with anti-Mi-2 sera. Such reactivity was previously detected (13), the significance of

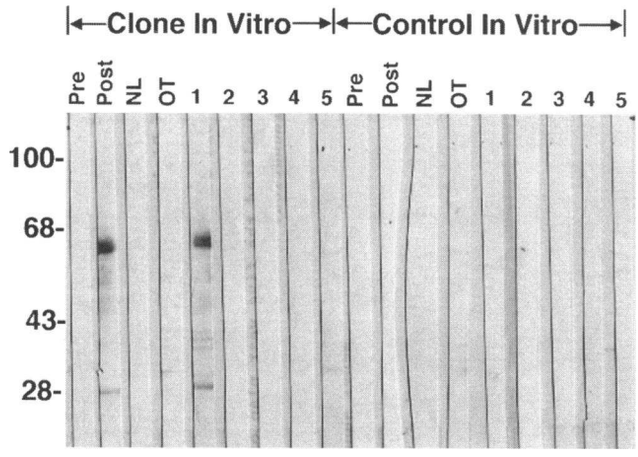

Figure 8. Immunoblot of in vitro transcribed and translated Mi-2 fragment: pGEM-3Z plasmid with clone C4 cDNA *(Clone)*, prepared by PCR with added start and stop codons, and the luciferase control *(Control)* were used as template in a coupled rabbit reticulocyte lysate transcription/translation system (Promega Corp.). 5 μl translation lysate was applied to each lane of the 10% SDS-PAGE, transferred to nitrocellulose, and blotted with the indicated sera (rabbit sera at 1:500, human sera at 1:200): *(Pre)* preimmune rabbit serum; *(Post)* postimmune serum; *(NL)* normal human serum; *(OT)* anti-Mi-2 negative anti-Ro/SSA-positive serum; *(1–5)* anti-Mi-2 sera: 1 = serum VA; 2–3 = other 240-kD blot positives; 4–5 = 240-kD blot negatives. The 60-kD protein stained by postimmune rabbit serum and patient serum VA, of predicted size for C4 cDNA product, corresponded in size to the radiolabeled translation product (not shown).

which will require further study. The affinity-purified antibodies studied here did not react with other Mi-2 components.

Epitopes. It is noteworthy that sera from all 44 patients reacted with the portion of the 240-kD protein encoded by the L4 cDNA, representing < 20% of the molecule. Furthermore, reaction with this portion was dependent on more than just the linear sequence for at least some of the sera, since they did not blot the in vitro–produced protein, even when they blotted the whole natural protein. This suggests that the immunoblot activity against the whole protein usually represents reaction with other portions of the protein not included in L4, and that the epitope(s) recognized by most sera in the plaque test were conformational, and either disrupted by the immunoblot procedure or not formed in vitro. These findings suggest a similarity in the response of different anti–Mi-2 patients, although serum VA showed significant additional reactivity.

The possibility of a predominant or universally shared epitope is consistent with a molecular mimicry hypothesis, although other concepts could also explain this. The factors which lead to production of autoantibodies to this protein in some DM patients, and not in other diseases, remain unclear.

Analysis of sequence. The cDNA of clone C4 is a portion of a larger cDNA, which appears to be 7.5 kb. It represents only 20% of the full length Mi-2 transcript, and 25% of a 240-kD protein, the average estimated relative size of the HeLa protein over a series of experiments. 7.5 kb is adequate to encode a 240-kD protein.

It was of interest that all four initial cDNA L clones were truncated within the coding region at identical points on both ends, despite coming from two different libraries. This suggests a specific reason for susceptibility of mRNA or cDNA to degradation at these points, or secondary structure resistant to denaturation. Attempted extension of the 3′ cDNA end by anchored PCR, using oligo dT and gene-specific primers (15), was unsuccessful.

Analysis of the predicted amino acid sequence encoded by clone C4 revealed several interesting features. Charged regions were found, which have been noted to be characteristic of autoantigens (26). Two were acidic regions, one of which was particularly rich in glutamic acid (aa 310–330) (Fig. 5 *A*). Similar acidic domains in GAL4 were felt to be important in transcriptional activation (27, 28). Two were basic regions, which would be expected to have a high affinity for nucleic acids (29). There were two proline-rich regions, which have been seen in proteins that bind single stranded nucleic acid, including proteins of Sm autoantigen (30, 31). Three regions of relative abundance of proline (P), glutamic acid (E), serine (S), threonine (T), and aspartic acid (D) (PEST or PEST/D regions) were noted (32) (Fig. 5 *A*). Such regions have been associated with proteins that undergo rapid intracellular degradation.

The encoded protein had four potential zinc finger motifs, although it is not known if these regions bind zinc or nucleic acids in this protein (33–35). The histidine-cystine pattern differs from that of TFIIIA Zn fingers and other recognized families. The EST-encoded homologous protein was of particular interest, since it was a different protein fragment with an almost identical Zn-finger region, that included two Zn-fingers and an NH_2 -terminal extension (Fig. 5 *B*). A 32 aa region of the EST was identical to Mi-2, and includes its entire NH_2-terminal Zn finger, plus the region between Zn fingers. The COOH-terminal Zn finger of the EST had the same form as that of Mi-2, with nearly complete identity in the putative Zn-binding regions and the same size loop. The limited reported sequence prevents more extensive analysis to determine its overall relatedness to Mi-2. The structure of the two sequential Zn fingers seen in Mi-2 and EST03869, as defined by the putative Zn-binding amino acids (C-X-X-C...C-X-X-C...H-X-X-C...C-X-X-C), was found in 14 other proteins, including the human proteins hrx, laminin a, laminin b1, and erb b-2 receptor, as well as the rat laminin b3 and neu, a papillomavirus protein, and others. The possible relationship of these proteins remains to be determined.

The presence of Zn fingers, proline-rich regions and charged regions suggests nucleic acid binding, raising the possibility of a role in DNA processing, regulation of transcription, or similar cellular processes, consistent with its nuclear localization. No nucleic acid has been found in association with Mi-2 antigen (9, 13), but this does not exclude nucleic acid binding. The fact that this protein is widely expressed and evolutionarily conserved (9) suggests that it is important in cellular function, but its role remains unknown, and the cDNA described here may help in determining its function.

Acknowledgments

Oligonucleotides were synthesized by the Molecular Biology Resource Facility of the William K. Warren Medical Research Institute at the University of Oklahoma Health Sciences Center. The authors wish to thank Drs. Morris Reichlin, Frank Arnett, Paul Plotz, Frederick Miller, Richard Leff, Jeffrey Callen, and Chester Oddis for referral of patient sera used in testing; Dr. Kazuko Itoh and Sally Blechinger for assistance with DNA purification and sequencing; Edward Trieu for expert technical assistance; Dr. Volker Brendel, Department of Mathematics, Stan-

ford University, Stanford, CA for allowing us to use the Statistical Analysis of Protein Sequences program; and Lester Cummings, Ross Thanscheidt, and Tom Barton of Management Information Systems of Oklahoma Medical Research Foundation for their invaluable computer and programming assistance.

This work was supported by grants AI27181 and AR32214 from the National Institutes of Health and by a grant from the Oklahoma Center for the Advancement of Science and Technology and by Veterans Affairs Medical Research Funds. Dr. Nilasena was a fellow of the Presbyterian Health Foundation during this study.

References

1. Targoff, I. N. 1992. Autoantibodies in polymyositis. *Rheum. Dis. Clin. North Am.* 18:455–482.

2. Targoff, I. N. 1993. Humoral immunity in polymyositis/dermatomyositis. *J. Invest. Dermatol.* 100:116S–123S.

3. Nishikai, M. and M. Reichlin. 1980. Heterogeneity of precipitating antibodies in polymyositis and dermatomyositis: characterization of the Jo-1 antibody system. *Arthritis Rheum.* 23:881–888.

4. Mathews, M. B. and R. M. Bernstein. 1983. Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity. *Nature (Lond.).* 304:177–179.

5. Bunn, C. C., R. M. Bernstein, and M. B. Mathews. 1986. Autoantibodies against alanyl-tRNA synthetase and $tRNA^{ala}$ coexist and are associated with myositis. *J. Exp. Med.* 163:1281–1291.

6. Mathews, M. B., M. Reichlin, G. R. V. Hughes, and R. M. Bernstein. 1984. Anti-threonyl-tRNA synthetase, a second myositis-related autoantibody. *J. Exp. Med.* 160:420–434.

7. Targoff, I. N. 1990. Autoantibodies to aminoacyl-transfer RNA synthetases for isoleucine and glycine: Two additional synthetases are antigenic in myositis. *J. Immunol.* 144:1737–1743.

8. Targoff, I. N., A. E. Johnson, and F. W. Miller. 1990. Antibody to signal recognition particle in polymyositis. *Arthritis Rheum.* 33:1361–1370.

9. Targoff, I. N., and M. Reichlin. 1985. The association between Mi-2 antibodies and dermatomyositis. *Arthritis Rheum.* 28:796–803.

10. Love, L. A., R. L. Leff, D. D. Fraser, I. N. Targoff, M. C. Dalakas, P. H. Plotz, and F. W. Miller. 1991. A new approach to the classification of idiopathic inflammatory myopathy: Myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine (Baltimore).* 70:360–374.

11. Targoff, I. N., D. S. Nilasena, E. P. Trieu, F. C. Arnett, L. M. Pachman, J. P. Callen, and F. W. Miller. 1990. Clinical features and immunologic testing of patients with anti-Mi-2 antibodies. *Arthritis Rheum.* 33:S72*a*. (Abstr.)

12. Rider, L. G., F. W. Miller, I. N. Targoff, D. D. Sherry, E. Samayoa, M. Lindahl, M. H. Wener, L. M. Pachman, and P. H. Plotz. 1994. A broadened spectrum of juvenile myositis: myositis-specific autoantibodies in children. *Arthritis Rheum.* 37:1534–1538.

13. Nilasena, D. S., E. P. Trieu, and I. N. Targoff. 1995. Analysis of the Mi-2 autoantigen of dermatomyositis. *Arthritis Rheum.* 38:123–128.

14. Young, R. A., and R. W. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194–1198.

15. Ge, Q., M. B. Frank, C. A. O'Brien, and I. N. Targoff. 1992. Cloning of a complementary DNA coding for the 100-kD antigenic protein of the PM-Scl autoantigen. *J. Clin. Invest.* 90:559–570.

16. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY. 1.53–1.69 and 9.52–9.53.

17. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

18. Kraft, R., J. Tardiff, K. S. Krayter, and L. A. Leinwand. 1988. Using miniprep plasmid DNA for sequencing double stranded templates with Sequenase. *Biotechniques.* 6:544–546.

19. Altschul, F. S., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tools. *J. Mol. Biol.* 215:403–410.

20. Brendel, V., P. Bucher, I. Nourbakhsh, B. E. Blaisdell, and S. Karlin. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. USA.* 89:2002–2006.

21. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid-guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–159.

22. Davis, L. G., M. D. Dibner, and J. F. Battey. 1986. Basic Methods in Molecular Biology. Elsevier Science Publishing, New York. 84–87.

23. Targoff, I. N., E. P. Trieu, and F. W. Miller. 1993. Reaction of anti-OJ autoantibodies with components of the multi-enzyme complex of aminoacyl-tRNA synthetases in addition to isoleucyl-tRNA synthetase. *J. Clin. Invest.* 91:2556–2564.

24. Kozak, M. 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic. Acids Res.* 15:8125–8148.

25. Adams, M. D., A. R. Kerlavage, C. Fields, and J. C. Venter. 1993. 3400 expressed sequence tags identify diversity of transcripts from human brain. *Nat. Genet.* 4:256–267.

26. Brendel, V., J. Dohlman, B. E. Blaisdell, and S. Karlin. 1991. Very long charge runs in systemic lupus erythematosus-associated autoantigens. *Proc. Natl. Acad Sci. USA.* 88:1536–1540.

27. Ma, J., and M. Ptashne. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell.* 48:847–853.

28. Sigler, P. B. 1988. Acid blobs and negative noodles. *Nature (Lond.).* 333:210–212.

29. Rokeach, L. A., J. A. Haselby, and S. O. Hoch. 1988. Molecular cloning of a cDNA encoding the human Sm-D autoantigen. *Proc. Natl. Acad. Sci. USA.* 85:4832–4836.

30. McAllister, G., A. Roby-Shemkovitz, S. G. Amara, and M. R. Lerner. 1989. cDNA sequence of the rat U snRNP-associated protein N: description of a potential Sm epitope. *EMBO (Eur. Mol. Biol. Org.) J.* 8:1177–1181.

31. Ohosone, Y., T. Mimori, A. Griffith, M. Akizuki, M. Homma, J. Craft, and J. A. Hardin. 1989. Molecular cloning of cDNA encoding Sm autoantigen: derivation of a cDNA for a B polypeptide of the U series of small nuclear ribonucleoprotein particles. *Proc. Natl. Acad. Sci. USA.* 86:4249–4253.

32. Rogers, S., R. Wells, and M. Rechsteiner. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science (Wash. DC).* 234:364–368.

33. Berg, J. M. 1986. Potential metal-binding domains in nucleic acid binding proteins. *Science (Wash. DC).* 232:485–487.

34. Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell.* 51:1079–1090.

35. Schiff, L. A., M. L. Nibert, M. S. Co, E. G. Brown, and B. N. Fields. 1988. Distinct binding sites for zinc and double stranded RNA in the reovirus outer capsid protein s3. *Mol Cell Biol.* 8:273–283.