Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8492–8496, November 1989
Immunology

Determination of an epitope of the diffuse systemic sclerosis marker antigen DNA topoisomerase I: Sequence similarity with retroviral p30^{gag} protein suggests a possible cause for autoimmunity in systemic sclerosis

(molecular mimicry/scleroderma/myasthenia gravis)

G. G. MAUL*, S. A. JIMENEZ, E. RIGGS, AND D. ZIEMNICKA-KOTULA

The Wistar Institute of Anatomy and Biology, 36th and Spruce Streets, Philadelphia, PA 19104

Communicated by George B. Koelle, June 21, 1989

ABSTRACT **The possibility that viruses play a role in the etiology of various autoimmune diseases has been proposed. One approach to the search for these agents involves identifying potential crossreactive epitopes in viruses that infect cells of the immune system or of the target tissues. Antibodies to DNA topoisomerase I are the marker autoantibodies for diffuse systemic sclerosis. The major epitope of the antigen was therefore sought through cloning and sequencing of the cDNA for human topoisomerase I and eventually by the synthesis of the smallest possible peptide recognized by sera from patients with the diffuse form of systemic sclerosis. The antigenic 11-amino acid sequence contains 6 sequential amino acids that are identical to a sequence present in the group-specific antigen (p30^{gag}) of some mammalian retroviruses. This sequence is separated by only 1 amino acid from the retroviral epitope sequence that crossreacts with autoantibodies against the marker antigen for mixed connective-tissue disease and systemic lupus erythematosus, the 70-kDa polypeptide of U1 ribonucleoprotein particles. These findings suggest that a retroviral agent may be involved in the pathogenesis of systemic sclerosis and other connective tissue diseases and that antibodies to intracellular antigens are not involved in the pathogenesis of autoimmune disease but may be useful as footprints for tracking the potential etiological agent of autoimmune disease.**

Autoimmunity to intracellular antigens is considered to be an epiphenomenon. Direct pathological effects of the autoantibodies are unlikely, since it has not been conclusively demonstrated that immunoglobulins can penetrate into the cytoplasm of healthy, intact cells. Tissue damage in the autoimmune diseases is usually secondary as in the case of DNA/anti-DNA antibody complexes, which may cause glomerulonephritis in systemic lupus erythematosus (SLE) (1). Although the mechanisms responsible for autoimmunity are not known, it may have been induced by a viral agent in genetically predisposed individuals (2–4). Any virus capable of attacking the immune system or the target organ(s) of an autoimmune disease should be scrutinized for evidence of etiological involvement in the disease. In autoimmune diseases, sequence similarities between proteins of potential infectious agents and autoantigens, a phenomenon described as molecular mimicry (5), might be the only remaining footprint of a disease-inducing agent. Antibodies to viruses that crossreact with a cell surface antigen, such as the acetylcholine receptor in myasthenia gravis (poliovirus capsid protein VP2) or myelin basic protein (hepatitis B virus DNA polymerase), can easily be integrated in hypotheses that postulate a direct role of viruses in the etiology of autoimmune diseases (6). This is not the case in diseases in which autoantibodies are directed against intracellular autoantigens. However, Query and Keene (7) reported an amino acid sequence similarity and antibody crossreactivity of the U1 ribonucleoprotein (RNP) 70-kDa marker antigen for mixed connective-tissue disease (MCTD) and SLE with the p30^{gag} protein of certain mammalian retroviruses. This finding supported previous observations (8) that many patients with autoimmune diseases have circulating antibodies reactive with a retroviral gag protein.

In diffuse systemic sclerosis (dSSc) the most characteristic autoantibodies are directed against an intracellular antigen, DNA topoisomerase I (EC 5.99.1.2) (9–11). The high specificity of these autoantibodies for dSSc renders them diagnostically relevant and indicates that their production is unlikely to be due to random B-cell activation. We have cloned and sequenced the major epitope of the dSSc marker antigen topoisomerase I and found that it contains a 6-amino acid match with the p30^{gag} retroviral protein.† This observation provides the conceptual basis for the hypothesis that the induction of autoimmunity to intracellular antigen(s) is caused by the same agent that is capable of inducing the disease through infection of target cells or cells of the immune system.

MATERIALS AND METHODS

Screening of λgt11 Expression Libraries. Several human λgt11 expression libraries (human T-cell, human endothelial, and human placental) obtained from Clontech were screened for topoisomerase I with serum from one patient with dSSc containing a high titer of anti-topoisomerase I antibodies by a chromogenic assay developed previously (12). The reagents used were prepared according to the manufacturer's instructions (Vector Laboratories). Freshly prepared 4-chloro-1-naphthol and 0.02% hydrogen peroxide were used as substrates. Positive plaques were tested against four sera containing the same antibodies as the serum used for the initial screening as predetermined by immunoblotting and electrophoretically fractionated HEp-2 human epidermoid carcinoma cell extract. The positive clones were used to generate recombinant lysogens (13). These lysogens were induced to produce the insoluble β-galactosidase fusion protein, which was then isolated as a granule (14).

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

Abbreviations: MCTD, mixed connective-tissue disease; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; dSSc, diffuse systemic sclerosis; CREST, calcinosis, Raynaud phenomenon, esophageal dysmobility, sclerodactyly, and telangiectasia syndrome; RNP, ribonucleoprotein.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M27913).

Verification of the Cloned cDNA. Several methods were used to verify that the cloned cDNAs encoded topoisomerase I. (*i*) Immunoblotting of the fusion protein against Scl-70 (scleroderma 70-kDa antigen) reference serum obtained from the Centers for Disease Control (Atlanta, GA). For this purpose, total HEp-2 cell extract or the isolated fusion protein was subjected to 10% polyacrylamide slab gel electrophoresis in the presence of sodium dodecyl sulfate (15) and electroblotted according to Towbin *et al.* (16). Strips (3 mm) were cut from the nitrocellulose sheet, placed in grooves cut into Lucite sheets, blocked with 1% goat serum and 1% Tween 40 for 30 min, and probed with human serum diluted 1:100. The reporter system was the same as for immunofluorescence. (*ii*) Immunofluorescence assay to determine the intracellular localization of antigens recognized by antibodies reactive with the fusion protein, which was previously affinity-purified by the method of Smith and Fisher (17). (*iii*) Immunoblotting to determine reactivity of antibodies that were affinity-purified with topoisomerase I isolated from electrophoretically fractionated HEp-2 cell extract. Since the serum autoantibodies were present at very high titer, an equal-sized area of the nitrocellulose strip close to the antigen-containing location was also incubated to test whether a substantial amount of antibody was nonspecifically bound to the membrane. (*iv*) Immunoblotting of 130 sera from patients with systemic sclerosis (SSc) to correlate their Scl-70 positivity with their reactivity with the fusion protein. The size of topoisomerase I mRNA was determined using cDNA from positive clones and total HeLa cell RNA (18). DNA·DNA hybridization (19) was used to determine whether different-sized cDNA fragments from clones reacting with the same serum were overlapping or different. (*v*) Sequencing the smallest epitope containing a topoisomerase I cDNA insert by the primer-extension method (20) for comparison with the sequences of topoisomerase I clones from two yeast species (21, 22).

Immunofluorescence and Immunoblotting. Nuclear antigen localization was determined on fixed HEp-2 cells and immunoblots prepared and developed as described (10). Total HEp-2 cell Western blots were obtained from BioDiagnostics (Wynnewood, PA) to test for U1 RNP and topoisomerase I autoantibodies. All sera were used at a 1:100 dilution in phosphate-buffered saline for 1 hr at room temperature (23°C) and processed with the biotin–avidin system according to the manufacturer's recommendations (Vector Laboratories), using horseradish peroxidase as the reporter molecule for the immunoblots and fluorescein-labeled avidin for the fluorescence microscopy. After a final wash the DNA of the HEp-2 cells was stained for 3 min with Hoechst 33258 (0.5 μg/ml). Cells were observed with a Nikon Omniphot microscope and the fluorescence pattern was recorded.

Computer Analysis of Amino Acid Sequences and Preparation of Synthetic Peptides. Computer analysis of the shortest epitope-containing sequence was used to determine hydrophilicity and highest surface probability. Overlapping 30-amino acid sequences were used to screen for significant matches with the total GenBank data base (release 56) translated to amino acid sequences. Two oligopeptides with high surface probability and flanking sequences with high amphipathic α-helical scores (23) were synthesized on an Applied Biosystems model 430A and were used as antigen (5 μg per well) in an ELISA-based test. These peptides (peptides A and B) represent the topoisomerase I sequences GVPIEKIYNKTQREKFA and KLEVQATDREENKQIALGTS. An additional peptide (peptide C), in which the retroviral amino acid match was left intact but the sequence of peptide A was changed to WGVPIEKIYNKTSGRSLPW, resulting in a more hydrophobic carboxyl end, was also synthesized and tested.

RESULTS

By screening human λgt11 libraries with sera containing antibodies to topoisomerase I, several positive clones were identified. Removal of the cDNA from the phage by *Eco*RI digestion resulted in inserts of length 1.1–2.8 kilobases. All inserts cross-hybridized. The β-galactosidase fusion protein reacted with the reference serum and with affinity-purified anti-topoisomerase I antibodies (from the 105-kDa band of total HEp-2 cell extract). Also, immunoglobulin that was affinity-purified by using the fusion protein as ligand gave the specific topoisomerase I immunofluorescence pattern described previously (10). Of the Scl-70-positive sera (by double diffusion) of all available sera from dSSc patients, 83% (71/86) of those that were tested with the fusion protein after electroblotting reacted strongly with the fusion protein (data not shown). In addition, a few sera negative for Scl-70 were positive with the fusion protein. None of the sera from 100 normal individuals and 100 SLE patients reacted with the fusion protein. The cDNA clones hybridized with a 4.0-kilobase mRNA (data not shown), indicating that a substantial untranslated nucleotide sequence must exist, since topoisomerase I is 105 kDa in size. The fusion protein did not adsorb out anti-topoisomerase I activity, indicating that other epitopes were present on the 105-kDa protein.

To determine the reactive epitopes on the β-galactosidase–topoisomerase I fusion molecule, we selected the shortest (1.1 kb) cDNA insert for sequencing. This analysis revealed that the largest open reading frame was only 327 base pairs followed by a long untranslated sequence (Fig. 1) of 837 base pairs. This effectively reduces to 109 the number of amino acids containing the epitope. Strong evidence that we have cloned topoisomerase I comes from the comparison of the deduced amino acid sequence with those published for those of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (21, 22) (Fig. 2). Alignment for best fit reveals long sequences of exact matches and conservative changes, indicating that these proteins are related and conserved over extremely large evolutionary distances.

The deduced amino acid sequence was analyzed for several parameters that might serve to predict the localization of putative antigenic sites or T-cell epitopes. Two areas of high surface probability were found, one of them also containing highly significant amphipathic α-helices (amphipathic score of 39; significance begins with a score of 4; ref. 23). We therefore focused on this area and searched the amino acid translation of the GenBank library for sequence similarities (Fig. 3). Sequence matches were found with several mammalian $p30^{gag}$ proteins and with the acetylcholine receptor, the marker autoantigen for myasthenia gravis. Comparison of the sequence similarity of topoisomerase I with the gag proteins shows that the 6-amino acid match seen in feline sarcoma virus is juxtaposed to one that has been proven to be antigenic and crossreacting with the U1 RNP 70-kDa protein, a marker antigen in MCTD and SLE. We therefore synthesized oligopeptides containing the two amino acid sequences most likely to be the epitopes in the 109-amino acid-long topoisomerase I sequence reactive with topoisomerase I-positive dSSc sera. When testing both peptides in an ELISA-based system using serum dilutions of 1:2000, we found that most topoisomerase I-positive sera reacted much more strongly with peptide A than control sera (Fig. 4), indicating that the synthetic peptide A contained the reactive epitope of the β-galactosidase–topoisomerase I fusion protein.

To determine whether the sequence match of peptide A with $p30^{gag}$ contained the site reactive with topoisomerase I-positive sera, a peptide was synthesized that had the carboxyl-terminal 6 amino acids changed so that they were more hydrophobic (peptide C). The same topoisomerase I-positive sera reacted even more strongly with peptide C

```
CAGCTAGCAGATGCCCGGAGAGACCTGAAAAGTGCTAAGGCTGATGCCAAGGTCATGAAG
Q  L  A  D  A  R  R  D  L  K  S  A  K  A  D  A  K  V  M  K
GATGCAAAGACGAAGAAGGTAGTAGAGTCAAAGAAGAAGGCTGTTCAGAGACTGGAGGAA
D  A  K  T  K  K  V  V  E  S  K  K  K  A  V  Q  R  L  E  E
CAGTTGATGAAGCTGGAAGTTCAAGCCACAGACCGAGAGGAAAATAAACAGATTGCCCTG
Q  L  M  K  L  E  V  Q  A  T  D  R  E  E  N  K  Q  I  A  L
GGAACCTCCAAACTCAATTATCTGGACCCTAGGATCACAGTGGCTTGGTGCAAGAAGTGG
G  T  S  K  L  N  Y  L  D  P  R  I  T  V  A  W  C  K  K  W
GGTGTCCCAATTGAGAAGATTTACAACAAAACCCAGCGGGAGAAGTTTGCCTGGGCCATT
G  V  P  I  E  K  I  Y  N  K  T  Q  R  E  K  F  A  W  A  I
GACATGGCTGATGAAGACTATGAGTTTTAGCCAGTCTCAAGAGGCAGAGTTCTGTGAAGA
D  M  A  D  E  D  Y  E  F  *
GGAACAGTGTGGTTTGGGAAAGATGGATAAACTGAGCCTCACTTGCCCTCGTGCCTGGGG
GAGAGAGGCAGCAAGTCTTAACAAACCAACATCTTTGTGAAAAGATAAACCTGGAGATAT
TATAAGGGAGAGCTGAGCCAGTTGTCCTATGGACAACTTATTTAAAAATATTTCAGATAT
CAAAATTCTAGCTGTATGATTTGTTTTGAATTTTGTTTTTATTTTCAAGAGGGCAAGTGG
ATGGGAATTTGTCAGCGTTCTACCAGGCAAATTCACTGTTTCACTGAAATGTTTGGATTC
TCTTAGCTACTGTATGCAAAGTCCGATTATATTGGTGCGTTTTTACAGTTAGGGTTTTGC
AATAACTTCTATATTTTAATAGAAATAAATTCCTAAACTCCCTTCCCTCTCTCCCATTTC
AGGAATTTAAAATTAAGTAGAACAAAAAACCCAGCGCACCTGTTAGAGTCGTCACTCTCT
ATTGTCATGGGGATCAATTTTCATTAAACTTGAAGCAGTCGTGGCTTTGGCAGTGTTTTG
GTTCAGACACCTGTTCACAGAAAAAGCATGATGGGAAAATATTTCCTGACTTGAGTGTTC
CTTTTTAAATGTGAATTTTTATTTCTTTTTAATTATTTTAAAATATTTAAACCTTTTTCT
TGATCTTAAAGATCGTGTAGATTGGGGTTGGGGAGGGATGAAGGGCGAGTGAATCTAAGG
ATAATGAAATAATCAGTGACTGAAACCATTTTCCCATCATCCTTTGTTCTGAGCATTCGC
TGTACCCTTTAAGATG
```

FIG. 1. Nucleotide sequence and deduced amino acid sequence of the shortest cDNA retaining the reactive epitope of the fusion protein. The 6-amino acid sequence that matches the sequence of $p30^{gag}$ retroviral protein is underlined.

than with peptide A (Fig. 4). This indicates that the epitope is contained in the amino-terminal segment of the peptide.

Assuming that the same type of retrovirus is responsible for autoimmune diseases like MCTD and SSc, that the epitopes are physically very close or overlapping on $p30^{gag}$, and that the antibodies are irrelevant to the etiology of the disease, one should predict that in each of the diseases the same type of autoantibodies can but do not necessarily appear. This prediction is borne out. Fig. 5 shows immunoblots of sera from MCTD patients (lanes A, E, and I). The antibodies recognized not only the 70-kDa antigen and other U1 RNP antigens like Sm (lanes A, E, and I) but also topoisomerase I (lanes A and I). A dSSc patient's serum recognized the 70-kDa antigen of U1 RNP but not topoisomerase I. As expected, serum from an Scl-70-positive dSSc patient recognized topoisomerase I (lane F). Three dSSc patients' sera recognized both antigens with the retroviral amino acid match (lanes C, D, and H). A confirmed CREST patient had topoisomerase I antibodies (lane G). However, the proportion of patients in the different disease groups having either one of those two autoantibodies are vastly different and many patients do not have either one of them (data not shown from immunoblot analysis of 25 MCTD patients, 180 dSSc patients, and 117 CREST patients).

DISCUSSION

Our conclusion that we have cloned a human topoisomerase I fragment comes from the reactivity of 83% of topoisomerase I-positive sera with the β-galactosidase fusion protein, from immunofluorescence localization of affinity-purified IgG, and from the strong sequence homology with the published topoisomerase I sequences from *S. cerevisiae* and the fission yeast *Sch. pombe* (20, 21). The human topoisomerase I sequence recently established (29) corresponds exactly with the partial one reported here. In all three cases, it is the carboxyl end of the molecule that shows the highest sequence homology. This finding also indicates that if the major epitope of the dSSc marker antigen is at the carboxyl end of topoisomerase I, then the 35-kDa peptide that is proteolytically cleaved from the 105-kDa topoisomerase I molecule (10, 11) to yield the 70-kDa autoantigen (24), must be at the amino terminus. Whether pathogenic or nonpathogenic yeasts could induce the dSSc-specific antibodies would have clinical relevance only if the antibodies were pathogenic—i.e., if the antigen were present on the surface of the target cells. This is clearly not the case for topoisomerase I.

Epitope determination in autoimmune diseases may lead to new concepts of disease induction and/or separate the phenomenon of autoimmunity from the process of disease in-

```
EKVDEKKQEFEKELKTGEVELKSSWNSVEKIKAQVEKLEQRIQTSSIQLKDKEEN
.  : ...          .          ::  :  :..::. .    .:  :.::: 
QLADARRDLKSAKADAKVMKDAKTKKVVESKKKAVQRLEEQLMKLEVQATDREEN
 :  :    :   :. :  :    .  .:   : . .:.: .  .  :  :..::
RLKAADELKKALDAELKSKKVDPGRSSMEQLEKRLNKLNERINVMRTQMIDKDEN


SQVSLGTSKINYIDPRLSVVFCKKYDVPIEKIFTKTLREKFKWAIESVDENWRF
 :. ::::::.::.:::. : .::::.:::::::.:: :::: ::::  ::.. :
KQIALGTSKLNYLDPRITVAWCKKWGVPIEKIYNKTQREKFAWAIDMADEDYEF
:  :::::::.::.:::.:  ..:. .::::::.. :: :.:: :: :    :. .
KTTALGTSKINYIDPRLTYSFSKREDVPIEKLFSKTIRDKFNWAAD.TPPDWKW
```

FIG. 2. Best fit among the deduced amino acid sequences of topoisomerase I from *S. cerevisiae* (top line), human (middle line), and *Sch. pombe* (bottom line). Strong homology exists at the carboxyl terminus. Identities and conservative changes are indicated by double and single dots, respectively.

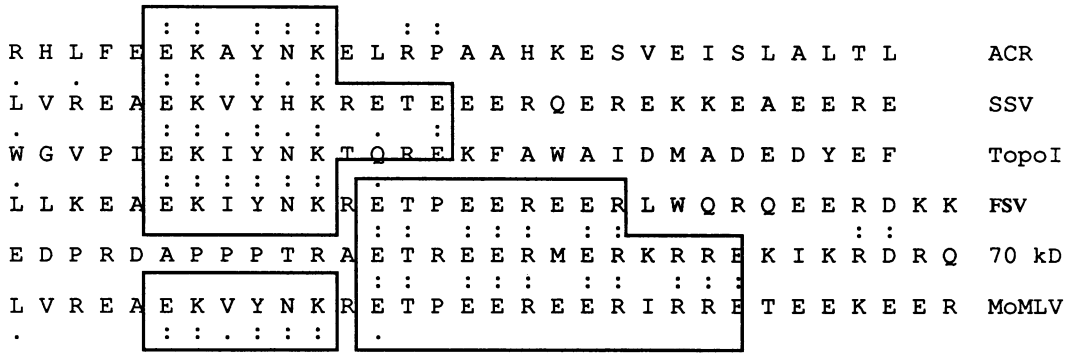

FIG. 3. Comparison of sequence similarities between three mammalian retroviral group-specific ($p30^{gag}$) proteins and three marker autoantigens, antibodies against which are markers for certain autoimmune diseases. Alignment of the single-letter amino acid code was made with the program for best fit. The areas of sequence matches for topoisomerase I with some of the retroviral $p30^{gag}$ sequences are boxed, as are those for the 70-kDa antigen of U1 RNP, demonstrating their close proximity in the aligned sequences. Double dots indicate exact matches, and single dots conservative changes. Dots in the top row (ACR) relate to the matches with the retroviral sequences. Dots in the bottom sequence (MoMLV) correlate with the topoisomerase I sequence. ACR, acetylcholine receptor (marker autoantigen for myasthenia gravis); SSV, simian sarcoma virus; TopoI, DNA topoisomerase I; FSV, feline sarcoma virus; 70 kD, 70-kDa U1 RNP antigen (autoantigen for MCTD and SLE); MoMLV, Moloney murine leukemia virus.

duction. The 109-amino acid sequence of the cloned topoisomerase I fragment effectively narrows the area containing the epitope recognized by autoantibodies to 10% of the 105-kDa molecule. Two potential antigenic determinants of this fragment are indicated by high surface probability (25) of the predicted amino acid sequence. One of these sequences was found to have a 6-amino acid match with some mammalian retrovirus group-specific antigens ($p30^{gag}$). The probability of encountering a particular sequence in unrelated proteins is the product of the probabilities to have the particular amino acids at any specific positions (26). If one uses the normalized frequencies for amino acids for unrelated proteins, it can be calculated that a particular hexamer in topoisomerase I should have a low probability (1.7×10^{-8}) of being present in an unrelated protein. The $p30^{gag}$ proteins in which this sequence was found are not related to topoisomerase I. Further, the probability that an additional autoantigen has sequence similarity with $p30^{gag}$ is in the same magnitude. The probability that it happens with two proteins is then the product of these probabilities. The number of antigens that are markers for autoimmune diseases is very low and the discrepancy of finding an epitope-sized amino acid match easily apparent. In the computer search, we also found a 5-amino acid match with the γ chain of the acetylcholine receptor, part of the market antigen of another autoimmune disease, myasthenia gravis. This match is nearly identical to the topoisomerase I amino acid match. The probability of having those 5 specific amino acids at a particular position is $\approx 4 \times 10^{-7}$.

A biological relevance of the 6-amino acid sequence match between the dSSc-specific autoantigen and a retroviral anti-

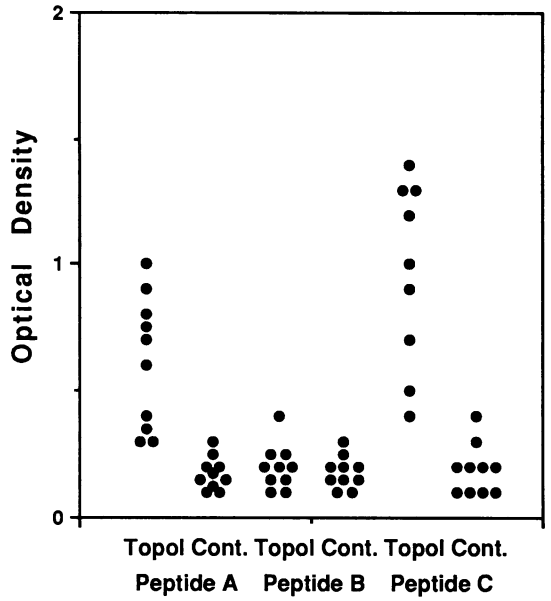

FIG. 4. Peptides were tested in an ELISA system. Peptides A and B represent topoisomerase I segments with a high antigenic index. Peptide C is a modified peptide A retaining the $p30^{gag}$ sequence match. Each peptide was tested with 10 topoisomerase I-positive (TopoI) sera and 10 control (Cont.) sera.

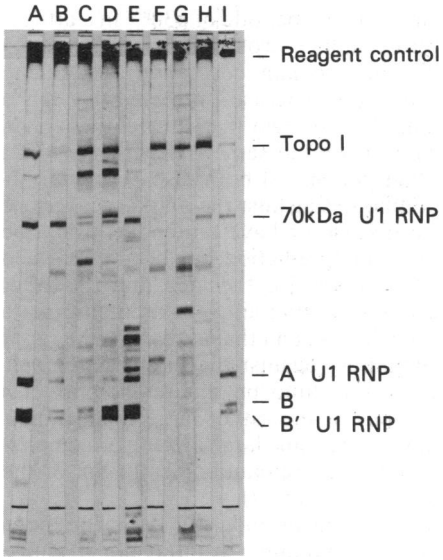

FIG. 5. Immunoblot analysis of selected sera. Sera from MCTD, dSSc, and CREST (calcinosis, Raynaud phenomenon, esophageal dysmobility, sclerodactyly, and telangiectasia syndrome) patients were tested using the peroxidase reporter system. Lanes A and I (MCTD) and C, D, and H (dSSc) show that serum from patients of each disease group can recognize both marker antigens at the same time. Lane B shows the strongest autoantibody to be directed against the 70-kDa antigen in this dSSc patient, and lane G indicates the presence of topoisomerase I (Topo I) autoantibodies in a male CREST patient. Except in lane H (dSSc), all other patients with the 70-kDa autoantibodies also had autoantibodies against other U1 RNP antigens like Sm and A, although in varying amounts. Lane E (MCTD) has all U1 RNP antibodies and lane F (dSSc) has strong topoisomerase I autoantibodies but no U1 RNP antibodies, as expected for these disease groups.

gen can be envisioned only if the sequence matches are situated in an area of antigenicity in both proteins. In both antigens, the amino acid match is located in an area that has a strong amphipathic α-helical index, indicating a high chance that this area contains potential T-cell epitopes (23). The amphipathic score for several of the $p30^{gag}$ sequences is >50 by the algorithm developed by Margalit *et al.* (23). For the topoisomerase I sequence the amphipathic score is 39, where >4 is the level at which significance begins. Whether the viral and topoisomerase I amphipathic α-helical regions are T-cell epitopes may be shown by examining the proliferative capacity of lymphocytes from scleroderma patients in the presence of the peptides encompassing the amphipathic sequences.

Testing the two sequences with the highest surface probability for reactivity with topoisomerase I-positive sera demonstrated that the one containing the 6-amino acid match contained the B-cell epitope with which dSSc sera reacted. Changing 6 amino acids of this peptide but leaving the $p30^{gag}$ 6-amino acid match intact retained reactivity. This test reduced the essential epitope size to 11 amino acids. Because the antibodies block enzyme activity (10, 11), we expected the epitope to be at the active site of topoisomerase I (28), which centers around the tyrosine residue on the amino-terminal side of the epitope determined here.

From the work of Query and Keene (7), it is known that $p30^{gag}$ crossreacts with MCTD patient sera containing antibodies that recognize the 70-kDa RNP antigen. The 70-kDa epitope sequence, expected by sequence matches to be the equivalent of the $p30^{gag}$ epitope, is separated by only 1 amino acid from the 6-amino acid match of topoisomerase I with the same $p30^{gag}$ (Fig. 3). We assume this to be more than a remarkable coincidence. The amphipathic α-helical region (potential T-cell epitope) of the retroviral $p30^{gag}$ ends with the autoantigen sequence match—i.e., with the B-cell epitope. Presentation of such a peptide of $p30^{gag}$ in a different context could therefore induce crossreactivity to topoisomerase I, and B cells may remain constantly restimulated by the availability of topoisomerase I released through cell death.

The finding that the two marker epitopes for MCTD and dSSc are present on the same retroviral protein allows the prediction that patients of both disease groups should exist who have antibodies against the marker antigen of the other autoimmune disease or have antibodies against both these autoantigens. This prediction has been confirmed and provides further support for the idea that the autoantibodies against these intracellular antigens are not important for the etiology of the disease and that autoantibodies are restricted. Also, autoimmunity against the virus-related epitopes may have existed at some time, but immunodominance has shifted to other epitopes of the rather large topoisomerase I or to other proteins of the same large molecular complex (U1 RNP particle), resulting in autoantibodies to Sm or the A and C antigens of this complex (7).

Our findings could provide the basis for the conceptual resolution of autoimmunity to intracellular antigens whose autoantibodies are unlikely to be pathogenic. The disease-specific autoantibodies would be epiphenomena but represent footprints with which to identify the virus. The actual virus-induced damage may be mediated by retrovirally transformed T cells or target cells (fibroblasts, endothelial cells) presenting retroviral antigen, such as $p30^{gag}$ and would therefore be independent of the humoral autoimmune response. Human retroviruses can transform and immortalize T cells (human T-cell lymphotropic virus type I) or lead to their destruction (human immunodeficiency virus) (see ref. 27 for review). Between these extremes, retroviral species might exist that immortalize and expand T-cell clones that are still responsive to normal control mechanisms and are not, therefore, capable of overwhelming the balance of the immune system in a manner analogous to that of Epstein–Barr virus-activated B cells. Such a scenario could result in slow cell-mediated damage induced at target cells having surface antigens with which the T cells are reactive. The hypothesis presented has several parts that are amenable to experimental verification and may yield clues regarding the pathogenesis of connective-tissue disease.

The technical assistance of Rena Hochenberg and Susanne Zürbig is acknowledged. This work was supported by a grant from the Scleroderma Research Foundation, a Sheryl N. Hirsch Award, and Grants AR38907, AM19106, and CA100815 from the National Institutes of Health.

1. Tan, E. M. (1982) *Adv. Immunol.* **33,** 167–240.
2. East, J., Prosser, P. R., Holborow, E. J. & Jaquet, H. (1967) *Lancet* **i,** 755–757.
3. Fujinami, R. & Oldstone, M. B. A. (1985) *Science* **230,** 1043–1045.
4. Kinsher, K. & Cunningham, W. W. (1985) *Science* **227,** 413–415.
5. Oldstone, M. B. A. (1987) *Cell* **50,** 819–820.
6. Fujinami, R. S., Oldstone, M. B. A., Wroblewska, A., Frankel, M. E. & Koprowski, H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2346–2350.
7. Query, C. C. & Keene, J. D. (1987) *Cell* **51,** 211–220.
8. Rucheton, M., Graafland, H., Fenton, H., Ursule, L., Ferrier, P. & Larsen, C. J. (1985) *Virology* **144,** 468–480.
9. Guldner, H. H., Szostecki, C., Vosberg, H. P., Lakomek, H. J., Penner, E. & Bautz, F. A. (1986) *Chromosoma* **94,** 131.
10. Maul, G. G., French, B. T., van Venrooij, W. J. & Jimenez, S. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5145–5149.
11. Shero, J. H., Bordwell, B., Rothfield, N. F. & Earnshaw, W. C. (1986) *Science* **231,** 737–740.
12. French, T. T., Maul, H. M. & Maul, G. G. (1986) *Anal. Biochem.* **156,** 417–423.
13. Huynh, T. V., Young, R. A. & Davis, R. W. (1982) in *DNA Cloning*, ed. Glover, D. M. (IRL, Washington, DC), Vol. 1, pp. 49–78.
14. Bikel, I., Roberts, T. M., Bladon, M. T., Green, R., Amann, E. & Livingston, D. M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 906–910.
15. Laemmli, U. K. (1970) *J. Virol.* **50,** 884–894.
16. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4350–4354.
17. Smith, D. W. & Fisher, P. A. (1984) *J. Cell Biol.* **99,** 20–28.
18. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
19. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
21. Thrash, C., Bankier, A. T., Barrell, B. G. & Sternglanz, R. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4374–4378.
22. Uemura, T., Morino, K., Uzawa, S., Shiozaki, K. & Yanagida, M. (1987) *Nucleic Acids Res.* **15,** 9727–9739.
23. Margalit, H., Spouge, J. L., Cornette, J. L., Cease, K. B., Delisi, C. & Berzofsky, J. A. (1987) *J. Immunol.* **138,** 2213–2229.
24. Douvas, A. S., Achten, M. & Tan, E. M. (1979) *J. Biol. Chem.* **254,** 10514–10522.
25. Emini, E. A., Hughes, J. V., Perlow, D. S. & Boger, J. (1985) *J. Virol.* **55,** 836–839.
26. Dayhoff, M. O. & Orcut, D. C. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 2170–2174.
27. Wong-Staal, F. & Gallo, R. C. (1985) *Nature (London)* **317,** 395–403.
28. Eng, W., Pandit, S. D. & Sternglanz, R. (1989) *J. Biol. Chem.* **264,** 13373–13376.
29. D'Arpa, P., Machlin, P. S., Ratrie, H., Rothfield, N. F., Cleveland, D. W. & Earnshaw, W. C. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2543–2547.