Autoantigenicity of Ro/SSA antigen is related to a nucleocapsid protein of vesicular stomatitis virus

(anti-Ro/SSA/autoantibodies/systemic lupus erythematosus/Sjögren syndrome)

R. HAL SCOFIELD AND JOHN B. HARLEY[†]

Arthritis and Immunology Program, Oklahoma Medical Research Foundation; Department of Medicine, University of Oklahoma Health Sciences Center; and Veterans Administration Medical Center, Oklahoma City, OK 73104

Communicated by Donald F. Steiner, October 25, 1990

ABSTRACT The fine specificity of a reference human anti-Ro/SSA autoantibody-containing serum has been analyzed by using sequential overlapping octapeptides from the human 60-kDa Ro/SSA antigen. From preliminary data, the most antigenic octapeptide in the carboxyl-terminal 120 amino acid residues of Ro/SSA shares seven of eight amino acids with the nucleocapsid (N) protein from the Indiana serotype of vesicular stomatitis virus. A sequence comparison of Ro/SSA and N has unexpectedly revealed six small peptides shared by Ro/SSA and N. Fine specificity analysis with 531 octapeptides from the Ro/SSA sequence demonstrates that five of the six shared small peptides are bound by anti-Ro/SSA ($P = 0.00017$). A more powerful association is not present in 12,476 protein sequences similarly evaluated. In addition, the inclusion of single-residue gaps in Ro/SSA enlarges the sequence similarity of Ro/SSA and N for three of the five small shared antigenic peptides. This additional level of sequence similarity between Ro/SSA and N is unlikely to be the result of chance ($P < 0.0002$). While a number of models may explain these data, including independent immune responses to N and Ro/SSA, these results are also consistent with anti-Ro/SSA autoantibodies being the consequence of a specific viral exposure.

Autoantibodies binding to Ro/SSA antigen are often found in the sera of patients with systemic lupus erythematosus (SLE), Sjögren syndrome, subacute cutaneous lupus erythematosus, complete congenital heart block, and neonatal lupus dermatitis. Existing evidence strongly suggests that these autoantibodies have immunopathogenic potential (1, 2).

The presence of anti-Ro/SSA autoantibodies is associated with powerful genetic effects at the HLA-DQ and the T-cell receptor β chain genes (3–5). These results, along with the greater antigenicity of the human form of Ro/SSA (6), implicate the particular human Ro/SSA autoantigen, histocompatibility molecules, and T-cell receptors in the molecular control of this autoimmune response.

The function of Ro/SSA in cellular metabolism remains to be defined. Ro/SSA is an RNA-protein complex found in all mammalian cells. The four human Ro RNAs, termed hY RNAs, are uridine-rich products of RNA polymerase III (7). In addition to the well-known 60-kDa Ro/SSA peptide (8–10), at least three additional peptides are bound by some sera containing anti-Ro/SSA precipitins (11).

Evaluation of the fine specificity of autoantibodies binding to small peptides prepared from the 60-kDa Ro/SSA sequence has been initiated to explore the primary structures of other proteins for the association of their sequences with the autoantigenicity of Ro/SSA. The data show that the Ro/SSA peptides which share sequence identity with the nucleocapsid

(N) protein from the Indiana serotype of vesicular stomatitis virus tend to be bound by anti-Ro/SSA sera. These results constitute circumstantial evidence consistent with the possibility that the autoimmune response to Ro/SSA is related to the immune processing of an N protein.

MATERIALS AND METHODS

Materials. Sera were obtained under protocols approved for human investigation by the Oklahoma Medical Research Foundation. The sera containing anti-Ro/SSA autoantibodies formed precipitates against the Ro/SSA antigen in Ouchterlony immunodiffusion and were obtained from patients who fulfilled the 1982 criteria for SLE (12). Reagents for octapeptide synthesis were obtained from Cambridge Research Biochemicals (Cambridge, U.K.) and MilliGen (Burlington, MA). Goat anti-human γ -chain-specific IgG conjugated to alkaline phosphatase, *p*-nitrophenyl phosphate, Tween 20, and bovine serum albumin were obtained from Sigma.

Peptide Synthesis. Every possible overlapping octapeptide from the sequences of the 60-kDa Ro/SSA protein (9, 10) was prepared by standard methods on a solid support (13). The minor antigenic differences between the two 60-kDa Ro/SSA sequences were not substantive; therefore, results with the sequence obtained by Deutscher *et al.* (9) are presented.

Solid-Phase Assay for Antibody. IgG binding of a 1:100 serum dilution to Ro/SSA octapeptides was assessed by an enzyme-linked immunosorbent assay as previously described (8, 13) using an anti-human γ -chain-specific alkaline phosphatase conjugate and the *p*-nitrophenyl phosphate substrate. Peptides of known positive and negative reactivity were used to validate the assay and to standardize measurements.

Central Octapeptide Assignment. A central antigenic octapeptide was chosen for each isolated octapeptide or group of neighboring octapeptides which developed an A_{405} greater than an arbitrarily chosen threshold value (between 0.35 and 1.0). The central antigenic octapeptide of a group was chosen so as to maximize shared amino acid sequence with the near-neighbor octapeptides that were also bound by the serum and to minimize the number of selected antigenic octapeptides. When tetrapeptide or greater amino acid sequence identity was shared between the central reactive octapeptide and those near-neighbor octapeptides that also exceeded the threshold value only the most central antigenic octapeptide was arbitrarily chosen.

Abbreviations: SLE, systemic lupus erythematosus; N protein, nucleocapsid protein; NBRF, National Biomedical Research Foundation.

[†]To whom reprint requests should be addressed at: Oklahoma Medical Research Foundation, 825 Northeast Thirteenth Street, Oklahoma City, OK 73104.

Statistical Analysis. Sequence comparisons and data base searches have been performed by using programs modified from the University of Wisconsin Genetics Computer Group software (14). The association between Ro/SSA antigenicity and the sequence identity shared by Ro/SSA and a query protein has been estimated. First, contiguous octapeptides of the Ro/SSA sequence sharing at least four amino acids with a query protein sequence are collected into groups. The average size of these groups determines the average size of the remaining groups of octapeptides. Second, the location of the central antigenic peptide in the sequence is used to determine whether an octapeptide group is antigenic. The antigenicity alignment probability is then calculated from a one-tailed Fisher exact test.

The tetrapeptide identity score was derived as a measure of the density of shared tetrapeptides found for each 100 amino acids of the Ro/SSA and query sequences.

Probability strategies were developed to assess the sequence similarities found between Ro/SSA and N in the region just adjacent to the shared short antigenic sequences. Insertion of a single amino acid gap in the Ro/SSA sequence was permitted to expand the shared antigenic sequences. These near-neighbor similarities were then further extended by amino acid identity between Ro/SSA and N or by single-residue gaps in Ro/SSA until no additional amino acids could be added. Three methods were used to assess significance. First, the 35 proteins with the closest association between Ro/SSA antigenicity and tetrapeptide identity with Ro/SSA were individually evaluated for their capacity to enlarge the antigenic regions of sequence similarity by adding gaps in the Ro/SSA sequence. The total number of amino acids added for each protein was modeled to a Poisson distribution. Second, a simulation experiment was performed in which three regions of five amino acids and two regions of four amino acids were randomly paired within the Ro/SSA and N sequences. This random assignment was repeated 50 times. The amino acids adjacent to these regions were evaluated for sequence identity after inclusion of a single gap in the Ro/SSA sequence. The number of identical amino acids found was recorded for each set of random assignments and modeled to a Poisson distribution. Third, the tripeptides

shared by Ro/SSA and N were evaluated for their concordance with Ro/SSA antigenicity and for their capacity to be extended by single amino acids in the Ro/SSA sequence. To be considered antigenic all three amino acids of the tripeptide were required to be contained in a central antigenic octapeptide of the reference serum. Some contiguous areas of shared sequence much longer than six octapeptides were found as a result of overlapping groups of shared tripeptides. Each contiguous group was divided into $N \div 6$ groups of octapeptides with the remainder arbitrarily distributed, where N is number of contiguous octapeptides which have at least tripeptide identity. Probability of an association was estimated by a one-tailed Fisher exact test.

RESULTS

Fine Specificity of Anti-Ro/SSA. The antigenicity of 531 overlapping octapeptides of Ro/SSA was tested with an anti-Ro/SSA reference serum and compared with a control serum (Fig. 1). Clearly the anti-Ro/SSA reference serum bound the octapeptides to a much greater extent than did the control serum. Repeated analysis with the anti-Ro/SSA reference serum showed that while the relative binding to individual octapeptides varied the pattern of binding remained remarkably consistent.

Antigenicity of Ro/SSA and Sequence Identity with N Protein. This approach was extended from previous work in which the antigenicity of the carboxyl-terminal 13-kDa peptide fragment of the 60-kDa Ro/SSA peptide was evaluated (R.H.S., W. D. Dickey, K. W. Jackson, and J.B.H., unpublished data). The octapeptide beginning at amino acid residue 485, EYRKKMDI, had the greatest antigenicity with a reference anti-Ro/SSA serum in the carboxyl-terminal 13-kDa fragment of Ro/SSA. A search of the 12,476 protein sequences in Release 21 from June 1989 of the National Biomedical Research Foundation (NBRF) data bank found the sequence EYRKKLMD from the N protein of the Indiana serotype of vesicular stomatitis virus (15). The inclusion of a single amino acid residue gap (*) in the Ro/SSA sequence of EYRKK*MD produced a sequence strikingly similar to EYRKKLMD found in N (Fig. 2, Ro/SSA and VSV-IND#1).

FIG. 1. Binding of a human anti-Ro/SSA reference serum to overlapping octapeptides of Ro/SSA. Every possible octapeptide from the human 60-kDa Ro/SSA amino acid sequence (13) has been synthesized onto a solid-phase support. An enzyme-linked immunosorbent assay has been used to detect binding by human antibody. Octapeptide numbers indicate the sequence position of their amino-terminal amino acid. (A) Binding of a normal control serum. (B) Binding of the reference anti-Ro/SSA serum. (C) Central antigenic octapeptides are identified by broken lines. (D) Solid blocks indicate tetrapeptide and longer regions of the Ro/SSA sequence shared with the N protein of the Indiana serotype and average 5.5 octapeptides since there are three tetrapeptide and three pentapeptide identities. The remainder of the Ro/SSA sequence is arbitrarily partitioned into groups with an average size of 5.5 octapeptides (open blocks). All groups are continuous except the remainder octapeptides, which have been collected together from different parts of the sequence. (E) The central antigenic octapeptides from three additional anti-Ro/SSA precipitin-positive sera are indicated by ∇ .

Ro/SSA	<u>52</u>	<u>57</u>	<u>133 136</u>	<u>185 189</u>	<u>237 241</u>	<u>264 267</u>	<u>278 283</u>	<u>337</u>	<u>342</u>	<u>350 354</u>	<u>406 410</u>	<u>485</u>	<u>491</u>
	<u>K • L • GLEN</u>		<u>KDLK</u>	<u>LSHL • K</u>	<u>RTKD • E</u>	<u>KSKE</u>	<u>LTALL • R</u>	<u>RGKL • K • W</u>	<u>LKALD</u>		<u>TEKD • S</u>	<u>EYRKK • MD</u>	
VSV-IND#1						<u>KSKE</u>	<u>LTALL / R</u>	<u>RGKL / K / W</u>	<u>LKALD</u>			<u>EYRKK / MD</u>	
						<u>34 37</u>	<u>303 309</u>	<u>81 88</u>	<u>110 114</u>			<u>151 158</u>	
VSV-IND#2						<u>KSKE</u>	<u>LTALL / R</u>	<u>RGKL / K / W</u>				<u>EYRKK / MD</u>	
						<u>34 37</u>	<u>303 309</u>	<u>81 88</u>				<u>151 158</u>	
VSV-NJ	<u>K / L / GLEN</u>										<u>TEKD / S</u>	<u>YRKK</u>	
	<u>155</u>	<u>162</u>									<u>357 362</u>	<u>152 155</u>	
VSV-Chand.				<u>LSHL / K</u>	<u>RTKD / E</u>						<u>EKD • S</u>		
				<u>232 237</u>	<u>168 173</u>						<u>359 362</u>		
Rabies			<u>KDLK</u>										
			<u>34 37</u>										

FIG. 2. Shared small peptides between Ro/SSA and N proteins which overlap central antigenic peptides of Ro/SSA or which may be extended by single amino acid gaps (*) in the Ro/SSA sequence. Vesicular stomatitis virus Indiana serotype (VSV-IND#1) (15) and (VSV-IND#2) (16), New Jersey serotype (VSV-NJ) (17), Chandipura strain (VSV-Chand.) (18), and rabies virus (19) are presented. Amino acid residues are presented in the single-letter code and in their order on the 60-kDa Ro/SSA sequence; lowercase letters are used where gaps have been introduced in Ro/SSA. Amino acid residues in the Ro/SSA sequence which are part of a central antigenic octapeptide of the reference anti-Ro/SSA serum are indicated by dotted underlines. Shared sequences with no overlap with the central antigenic octapeptides or sequence extension by single amino acid gaps in Ro/SSA are not presented. DEMV is found at residues 417–420 in Ro/SSA and 256–259 in VSV-IND. DLLR is at 181–184 of Ro/SSA and 258–261 of VSV-Chand. The EVCR peptide from 117–120 of Ro/SSA is the only peptide shared with the Piry strain sequence at 258–261 (20). AAAM and DPDD are respectively found at 369–399 and 510–513 in Ro/SSA while at 75–78 and 66–69 in rabies.

This seven-of-eight amino acid sequence identity might seem to be very significant. The ordinarily applied assumptions of an equal partitioning of amino acids and a random distribution of amino acids within sequences would yield a probability of 6×10^{-9} for finding at least a seven-of-eight match. Even when allowance is made for the 3,318,616 octapeptides in the NBRF data base, this level of similarity appears to be unexpected. On the other hand, the assumptions underlying this analysis are testable.

Forty-eight protein sequences have been found to share a seven-of-eight amino acid sequence identity with Ro/SSA when the NBRF data base is searched with each of the 531 octapeptides of Ro/SSA. No greater similarity than shared seven of eight sequences has been found. The probability of a seven-of-eight amino acid identity from this data base with a randomly selected Ro/SSA octapeptide is 0.091, which is not sufficiently unlikely to reject the null hypothesis and suggests that the ordinarily applied assumptions may overestimate the significance of sequence similarities.

Nevertheless, if sequence similarity between N and Ro/SSA were reflected in one antigenic site, then other sites of Ro/SSA antigenicity might also be related by shared sequences. A tetrapeptide level of identity was initially chosen since this should be sufficiently infrequent to be important from the perspective of the number of comparisons made and yet be large enough to be of potential importance for antibody binding. Six shared short amino acid sequences of at least tetrapeptide length were found in both N and Ro/SSA [KSKE, LTALL, RGKL, KALD, DEMV, and EYRKK (see Fig. 2, VSV-IND#1)].

The central question, however, is whether structural similarity, in this case as measured by short peptide identity, is related to the antigenicity of Ro/SSA. The antigenicity alignment probability was developed to estimate the probability of an association between antigenicity and sequence identity. First, the central octapeptide of each potential epitope was chosen. Fourteen central antigenic octapeptides of Ro/SSA were identified from the binding of the reference serum (Fig. 1C). Second, the octapeptides of the Ro/SSA molecule were organized into groups based upon the presence or absence of tetrapeptide identity with Indiana strain N (shaded blocks in Fig. 1D). The remainder of the molecule was arbitrarily divided into groups of octapeptides of essentially the same average size. No Ro/SSA octapeptide of any of these latter groups shared tetrapeptide identity with the sequence of N. For the example of the reference anti-Ro/SSA serum, five of the six octapeptide groups of Ro/SSA

which shared tetrapeptide identity with N also contained a central antigenic peptide. On the other hand, only 9 of the 90 octapeptide groups which had no tetrapeptide identity with N contained a central antigenic peptide (Fig. 1C and D). The antigenicity alignment probability is 0.00017 for the association of the epitopes of Ro/SSA with tetrapeptide identity between Ro/SSA and N. These are, to our knowledge, the first data consistent with the hypothesis that different areas of linear sequence antigenicity might be related to multiple areas of short sequence identity between an autoantigen and a foreign protein.

Structural Relationships of Ro/SSA with NBRF Proteins. The linear sequence autoantigenicity of Ro/SSA is clearly associated with the primary structure of the N protein. The potential importance of this association depends upon how unexpected the relationship is. The 12,476 protein sequences in the NBRF database provide an opportunity to search for other protein sequences related to Ro/SSA autoantigenicity.

First, every protein in the NBRF database was assessed for shared tetrapeptides with the 60-kDa Ro/SSA sequence and the density of shared tetrapeptides was calculated as the tetrapeptide identity score. Sites of Ro/SSA antigenicity are not considered in this measure. There are 309 sequences or 2.5% of the data base which exceed the 0.401 tetrapeptide identity score achieved by N. This result means that the relationship of N to Ro/SSA antigenicity cannot simply be attributed to unusual primary structure similarity.

The second and most central issue is whether other proteins share structure with the antigenic sequences of Ro/SSA to a greater extent than that found with N. The antigenic alignment probability has been calculated for all 12,476 protein sequences of the NBRF database. Of these, 17 protein sequences have an antigenicity alignment probability of <0.01 . The two almost-identical N protein sequences from the Indiana serotype (NBRF access codes vhnvnn and vhnv4), however, have sequences shared with Ro/SSA which are the most closely related to Ro/SSA antigenicity ($P = 0.00017$ and $P = 0.0012$). Ribonucleoside diphosphate reductase from *Escherichia coli* (access code s00926) has the next most unlikely antigenicity alignment probability (0.0036). This is, however, over 20-fold more likely to have occurred by chance than is the relationship found with the best N protein sequence (vhnvnn, VSV-IND#1 in Fig. 2).

Tripeptide Identities Between Ro/SSA and N. The antigenic alignment probability for the reference anti-Ro/SSA serum has been revised for the case of tripeptide identity between N and Ro/SSA as a trial of methodologic robustness. Eight

of the 27 areas in the Ro/SSA sequence which share at least tripeptide identity with N contained central antigenic octapeptides, while only 6 of the 52 areas with no tripeptide identity were antigenic; therefore, the antigenic alignment probability for shared tripeptides is 0.048 with an odds ratio of 3.2. This is a much weaker association than found at the tetrapeptide level but nevertheless remains significant. Consequently, these data also support a specific autoimmune relationship between N and Ro/SSA at the level of shared tripeptides where the number of shared but antigenically irrelevant sequences may be substantially increased.

Additional Structural Similarity. The amino acid sequence of the Indiana strain N protein in the near neighborhood of the small peptides identical with Ro/SSA (Fig. 2) provides additional support to conclude that N from this virus strain has a special similarity to Ro/SSA. The inclusion of single amino acid gaps in the Ro/SSA sequence increases the sequence similarity between Indiana serotype N protein and Ro/SSA for three of the shared tetra- and pentapeptides, as presented in Fig. 2. In each instance the gap is found in the Ro/SSA sequence on the carboxyl side of an area of tetrapeptide or pentapeptide identity with Ro/SSA. Each of the three regions of Ro/SSA with this near-neighborhood similarity with Indiana strain N is antigenic. It is pertinent to note that single amino acid deletions do not always alter antigenicity (21). Four strategies were used to assess the importance of this observation.

The regions of tetrapeptide or greater identity of the six known N protein sequences were compared for their relationship to the antigenicity of Ro/SSA (Fig. 2). Both rabies (19) and the New Jersey serotype (17) shared only three tetrapeptides and no pentapeptides with Ro/SSA, while Chandipura (18) and Piry strains (20), respectively, shared four and one tetrapeptides. Two of the New Jersey serotype N protein sequences overlap with central antigenic peptide, while only one overlaps in each of the Chandipura and rabies N proteins (Fig. 2). Three additional amino acids may be added to the shared tetrapeptides of New Jersey serotype and Chandipura N proteins by allowing single amino acid gaps in the Ro/SSA sequence. None may be added to the shared rabies or Piry strain sequences by this rule. Only the Indiana serotype N proteins, however, achieve a meaningful antigenic alignment probability with the reference anti-Ro/SSA serum. Collectively, these observations suggest that the Indiana serotype N protein sequence is more closely related to the antigenicity of Ro/SSA than are the other known rhabdoviral N sequences.

The 35 protein sequences with the smallest antigenic alignment probability were assessed for their capacity to allow the addition of single amino acid residue gaps in Ro/SSA to extend shared sequences. For none of these could more than three amino acids be added. By a Poisson distribution, the five near-neighbor amino acids added between Ro/SSA and the Indiana serotype N protein (Fig. 2) are very unusual ($P = 0.00039$).

The likelihood that the shared near-neighbor amino acids discovered by allowing gaps in Ro/SSA was also assessed by a simulation experiment. The object was to control for amino acid composition of Ro/SSA and N as well as for the number of antigenic shared peptides. The sequence positions of two tetrapeptides and three pentapeptides were randomly and separately assigned to the Ro/SSA and N sequences. Amino acids in the near neighborhood of these tetra- and pentapeptides shared between the two proteins after allowing single amino acid residue gaps in the Ro/SSA sequence were then noted. This procedure was repeated 50 times. For no instance could more than three additional shared amino acids be added by allowing single residue gaps in Ro/SSA. By the Poisson distribution, the presence of five or more additional shared

amino acids as observed with Indiana strain N and Ro/SSA was very unexpected ($P = 0.00017$).

Finally, if there is a relationship between a primary sequence of N and the antigenicity of Ro/SSA, then one might expect the shared amino acids which are added by single residue gaps in Ro/SSA to be also related to the antigenic sites. To have a sufficiently large sample size, this hypothesis was tested at the level of tripeptides and longer sequences shared by N and Ro/SSA. Of the 38 pairs of shared tripeptides (or longer) there were five to which additional amino acids could be added by allowing a single residue gap in Ro/SSA. Four of these were in central antigenic octapeptides. Nine of the remaining 33 shared sequences were in central antigenic octapeptides of Ro/SSA. The resulting contingency table suggests an association between Ro/SSA antigenicity and the additional shared amino acids found with gaps in Ro/SSA ($P = 0.036$). Beyond the analysis of the antigenic alignment cited above, the magnitude of the sequence extensions is independent evidence that the sequence of Indiana serotype N is related to the autoantigenicity of Ro/SSA.

Other Anti-Ro/SSA Sera. The antigenicities of a total of six anti-Ro/SSA sera were evaluated for binding to the 531 octapeptides of Ro/SSA. Two sera which formed a precipitate with purified native 60-kDa Ro/SSA but did not convincingly bind 60-kDa Ro/SSA in immunoblots did not bind any of the Ro/SSA octapeptides above background. Presumably, the anti-Ro/SSA autoantibodies of these sera bound only conformational epitopes. Sera with similar behavior have previously been described in an immunoblot analysis (22) and in the experiments analyzing the fine specificity of the 13-kDa carboxyl terminus of Ro/SSA. The other sera bound an approximate subset of the epitopes bound by the reference serum as presented in Fig. 1. The positions of the central antigenic peptides for these sera are indicated in Fig. 1E. With the Indiana serotype N sequence (vhvnn), these sera produced antigenicity alignment probabilities of 0.0013, 0.0048, and 0.0081. None of these sera bound the octapeptides containing the shared pentapeptide LKALD. Since this pentapeptide was not shared with the second Indiana strain N protein sequence [VSV-IND#2 in Fig. 2 (vhvn4)] the antigenic alignment probabilities were higher with this N protein sequence: 0.00040, 0.0016, and 0.0027, respectively. While some anti-Ro/SSA sera do not appear to bind the linear sequence, in many others the sequential autoantigenicity of Ro/SSA is related to the sequence similarity of Ro/SSA with an N protein of vesicular stomatitis virus.

DISCUSSION

Autoantigenicity of Ro/SSA Peptides and N. The data and analysis present a strong association of the autoantigenicity of 60-kDa Ro/SSA with the primary structure of an N protein from the Indiana serotype of vesicular stomatitis virus. This conclusion is supported by two independent arguments. First, the antigenic sites in the linear sequence of Ro/SSA tend to share sequence identity with the N protein. Second, allowing single amino acid gaps in the Ro/SSA sequence at the shared antigenic sequences reveals additional unexpected sequence similarity.

The first alternative explanation is that though these relationships are interesting, they may hold no etiologic importance for the origin of anti-Ro/SSA autoantibodies or for the anti-Ro/SSA-associated diseases. If true, then the associations described are the consequence either of chance, no matter how unlikely, or of independently arising immune responses to N and Ro/SSA which have no influence or consequence upon one another. The second alternative explanation is that vesicular stomatitis virus N protein is important in the autoimmunity against Ro/SSA. Establishing

this hypothesis may prove to be very difficult, and the possibility that these results hold no etiologic significance must remain a serious consideration until direct evidence to the contrary is obtained.

By a corollary hypothesis the N protein of vesicular stomatitis virus would be related only to anti-Ro/SSA autoantibodies and not other autoantibodies. A general defect could permit disease expression, but different viruses or foreign antigens would serve as stimuli or immunogens for other observed autoantibodies and other autoimmune disease-related phenomena.

Vesicular Stomatitis Virus. Many features of vesicular stomatitis virus are, however, consistent with those of a potential etiologic agent of the anti-Ro/SSA-associated rheumatic autoimmune diseases. The vesicular stomatitis virus leader sequence binds La/SSB protein (23), autoantibodies against which are associated with anti-Ro/SSA in SLE. The earliest known effect of vesicular stomatitis virus infection upon cellular metabolism is the inhibition of U particle assembly (24). U particles are also common autoantigens in SLE (1). Vesicular stomatitis virus is thought to infect cells through nonspecific surface glycoproteins. Consequently, the virus has a tremendous host range, with isolates from plants growing well in mammalian tissue culture cells (25). While the tissue tropism of this virus may vary, a variety of human tissues would consequently be expected to support the viral infection, consistent with the clinical heterogeneity and unpredictable course of SLE. Vesicular stomatitis virus survives better in complement-deficient serum, and many complement component deficiency states predispose to SLE (26, 27). Patients with SLE produce high levels of interferon (28) which induces latency of vesicular stomatitis virus. Interferon inhibits the production of intact virions, which could explain the inability to find assembled viral particles in patient tissues (29). During latency viral antigens may be expressed on cell surfaces (30), which may also constitute a source of continuing antigenic stimulation. The N protein is immunogenic in whole virus infections of animals, resulting in cytotoxic T-lymphocyte responses directed against the N protein (31, 32). In addition, vesicular stomatitis virus, through its G protein, is known to be capable of generating HLA class II restricted T-lymphocyte responses (33). This is compatible with the known HLA class II and T-lymphocyte receptor associations with the anti-Ro/SSA autoimmune response (1, 3–5).

The major existing impediment to accepting vesicular stomatitis virus as being important in the anti-Ro/SSA response is its absence as an infection of livestock in Europe. Since anti-Ro/SSA and the associated human diseases occur in Europe as they do elsewhere, perhaps none of the strains of the Indiana serotype of vesicular stomatitis virus now available are good representatives for virus purportedly involved with the anti-Ro/SSA autoimmune response. Over 70 strains of rhabdoviruses are known, however, to infect vertebrates, and probably hundreds remain to be described (34, 35). So there is a reasonable possibility that a rhabdovirus exists which would be an even better explanation for the autoantigenicity of Ro/SSA than are those whose N proteins have been sequenced (14–20). Only the N proteins from the Indiana serotype (Fig. 2), however, achieve a significant antigenic alignment probability with Ro/SSA. This result suggests that future attention should be focused upon the subtypes and other close relatives of the Indiana serotype.

We are grateful for the encouragement of M. Reichlin, A. Horsfall, B. Neas, and W. Dickey; for the foresight and institutional planning of W. G. Thurman, J. Tang, and L. R. Cummings; for the expert technical assistance of D. Farris; for the computer programming assistance of G. Koelsch; and for the support of the March of Dimes Birth Defects Foundation (Grant 1-1109), the National Institutes of

Health (Grants AI24717, AR39577, and AI21568), and the Veterans Administration. R.H.S. is a recipient of a W. W. Rucks Fellowship and a Presbyterian-Harris Fellowship. This work is dedicated to the memory of James E. Allen.

1. Harley, J. B. & Gaither, K. K. (1988) *Rheum. Dis. Clin. N. Am.* **14**, 43–56.
2. Lee, L. A., Gaither, K. K., Coulter, S., Norris, D. A. & Harley, J. B. (1989) *J. Clin. Invest.* **83**, 1556–1562.
3. Harley, J. B., Sestak, A. S., Willis, L. G., Fu, S. M., Hansen, J. A. & Reichlin, M. (1989) *Arthritis Rheum.* **32**, 826–836.
4. Fujisaku, A., Frank, M. B., Neas, B., Reichlin, M. & Harley, J. B. (1990) *J. Clin. Invest.* **86**, 606–611.
5. Frank, M. B., McArthur, R., Harley, J. B. & Fujisaku, A. (1990) *J. Clin. Invest.* **85**, 33–39.
6. Reichlin, M., Rader, M. D. & Harley, J. B. (1989) *Clin. Exp. Immunol.* **76**, 373–377.
7. O'Brien, C. A. & Harley, J. B. (1990) *EMBO J.* **9**, 3683–3689.
8. Yamagata, H., Harley, J. B. & Reichlin, M. (1984) *J. Clin. Invest.* **74**, 625–633.
9. Deutscher, S. L., Harley, J. B. & Keene, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9479–9483.
10. Ben-Chetrit, E., Gandy, B. J., Tan, E. M. & Sullivan, K. F. (1989) *J. Clin. Invest.* **83**, 1284–1292.
11. McCauliffe, D. P., Luk, F., Lieu, T. S., Sanz, I., Hanke, J., Newkirk, M. M., Bachinski, L. L., Itoh, Y., Reichlin, M., Sontheimer, R. D. & Capra, J. D. (1990) *J. Clin. Invest.* **85**, 1379–1391.
12. Tan, E. M., Cohen, A. S., Fries, J. F., Masi, A. T., McShane, D. J., Rothfield, N. F., Schaller, J. G., Talal, N. & Winchester, R. J. (1982) *Arthritis Rheum.* **25**, 1271–1277.
13. Geysen, H. M., Rodda, S. J., Wason, T. J., Tribbick, G. & Schoofs, P. G. (1987) *J. Immunol. Methods* **102**, 259–274.
14. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
15. Gallione, C. J., Greene, J. R., Iverson, L. E. & Rose, J. K. (1981) *J. Virol.* **39**, 529–535.
16. DePolo, N. J., Giachetti, C. & Holland, J. J. (1987) *J. Virol.* **61**, 454–464.
17. Banarjee, A. K., Rhodes, D. P. & Gill, D. S. (1984) *Virology* **137**, 432–438.
18. Masters, P. S. & Banarjee, A. K. (1987) *Virology* **157**, 298–306.
19. Tordo, N., Poch, O. & Keith, G. (1986) *Nucleic Acids Res.* **14**, 2671–2683.
20. Crysler, J. G., Lee, P., Reinders, M. & Prevec, L. (1990) *J. Gen. Virol.* **71**, 2191–2194.
21. Appel, J. R., Pinilla, C., Niman, H. & Houghten, R. (1990) *J. Immunol.* **144**, 976–983.
22. James, J. A., Dickey, W. D., Fujisaku, A., O'Brien, C. A., Deutscher, S. L., Keene, J. D. & Harley, J. B. (1990) *Arthritis Rheum.* **33**, 102–106.
23. Kurilla, M. G. & Keene, J. D. (1983) *Cell* **34**, 837–845.
24. Crone, D. E. & Keene, J. D. (1989) *J. Virol.* **63**, 4172–4180.
25. Superti, F., Seganti, L., Ruggeri, F. M., Tinari, A., Donelli, G. & Orsi, N. (1987) *J. Gen. Virol.* **68**, 387–399.
26. Beebe, D. P. & Cooper, N. R. (1981) *J. Immunol.* **126**, 1562–1568.
27. Agnello, V. (1987) in *Systemic Lupus Erythematosus*, ed. Lahita, R. G. (Wiley, New York), pp. 565–589.
28. Prebie, O. T., Black, R. J., Friedman, R. M., Kippel, J. H. & Vilcek, J. (1982) *Science* **216**, 429–431.
29. Sekellick, M. J. & Marcus, P. I. (1980) in *Rhabdoviruses*, ed. Bishop, D. H. L. (CRC, Boca Raton, FL), Vol. 1, pp. 69–98.
30. Holland, J. J., Villarreal, L. P., Welsh, R. M., Oldstone, M. B. A., Kohne, D., Lazzarini, R. & Scolnick, E. (1976) *J. Gen. Virol.* **33**, 193–211.
31. Wagner, R. R. (1987) in *The Rhabdoviruses*, ed. Wagner, R. R. (Plenum, New York), pp. 9–74.
32. Puddington, L., Bevan, M. J., Rose, J. R. & Lefrancois, L. (1986) *J. Virol.* **60**, 707–717.
33. Browning, M. J., Huang, A. S. & Reiss, C. S. (1990) *J. Immunol.* **145**, 985–994.
34. Brown, F. & Crick, J. (1979) in *Rhabdoviruses*, ed. Bishop, D. H. L. (CRC, Boca Raton, FL), Vol. 1, pp. 1–22.
35. Wagner, R. R. (1990) in *Virology*, eds. Fields, B. N. & Knipe, D. M. (Raven, New York), 2nd Ed., pp. 867–881.