

CHROMSYM. 1335

REVERSED-PHASE LIQUID CHROMATOGRAPHY AND SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS CHARACTERISTICS OF A RECOMBINANT DNA DERIVED MALARIA ANTIGEN

K. BENEDEK*, B. HUGHES, M. B. SEAMAN and J. K. SWADESH*
Smith Kline & French Laboratories, King of Prussia, PA 19406-0939 (U.S.A.)

SUMMARY

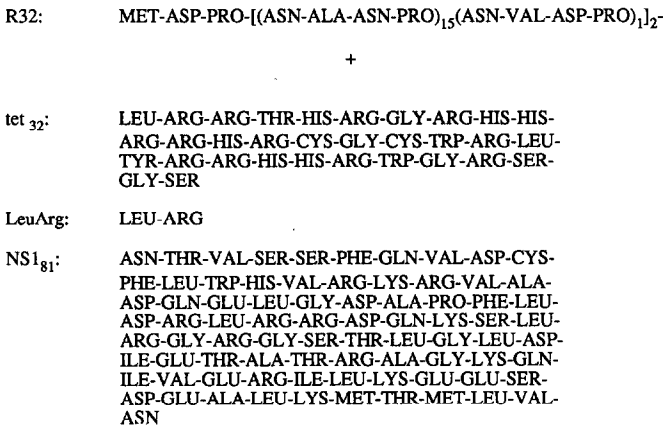
This paper presents results from analysis of a sample of SK&F 105154 (R32NSI₈₁), a malaria vaccine candidate produced in *Escherichia coli*, and discusses some analytical issues of general relevance to the characterization of such products derived from recombinant DNA technology. Anomalous migration and staining behavior were observed in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Reversed-phase liquid chromatography (RPLC) appeared to resolve four minor components from the principal band, but the minor peaks were found to consist of numerous components resolvable by SDS-PAGE. Western blotting visualized certain components that were not adequately stained by either Coomassie or silver stain. None of the techniques that were employed were individually adequate to characterize the sample, but, taken together, were adequate to characterize the sample and to identify one principal degradation pathway. Degradation within the NSI₈₁ region decreases the RPLC retention time, while degradation within the R32 segment increases the retention time.

INTRODUCTION

Immunization with irradiated sporozoites of *Plasmodium* is known to generate protective antibodies in animals and man¹⁻⁷. Several promising vaccines, based on the central domain of the circumsporozoite gene of *P. falciparum*, contain repeating sequence coupled in various ways to exogenous sequences⁸⁻¹¹; the exogenous sequences bind to some adjuvants, such as alumina, and may also serve as T-cell primers^{3,8,9}. The repeating sequence is Asn-Ala-Asn-Pro or the variant Asn-Val-Asp-Pro^{8,9}. Scheme 1 illustrates three of the vaccine candidate constructs, namely R32LeuArg, R32tet₃₂, and R32NSI₈₁; LeuArg indicates a leucine-arginine terminus, tet₃₂ indicates a 32-residue peptide derived from the tetracycline resistance gene, and NSI₈₁ indicates an 81-residue sequence derived from the N-terminus of the influenza A non-structural protein 1.

Electrophoretic analysis of constructs based on R32 is not routine⁹. The R32

* Present address: Centocor, Malvern, PA 19355, U.S.A.



Scheme 1. Primary structures of different recombinant *P. falciparum* CS malaria antigens. The constructs consists of two portions; *i.e.* R32, the antigenic core built from the repeating sequence of Asn-Ala-Asn-Pro tetrapeptide, and the tail sections tet₃₂, LeuArg and NS1₈₁.

segment, if not conjugated to an exogenous sequence is not appreciably stained by such common electrophoretic stains as Coomassie Brilliant Blue and/or glutaraldehyde-enhanced silver stain. It can be visualized by Western (antigenic) blotting using a mixture of R32tet₃₂ derived monoclonal antibodies as a broad, complex band at very high apparent molecular weight. R32tet₃₂ exhibits loading-dependent migration suggestive of aggregation¹¹ in the standard Laemmli system¹². Even at very low loadings, where a single band is observed, the apparent molecular weight of the principal band is much higher than the calculated based upon the known structure of this polypeptide. R32NS1₈₁ exhibits no evidence of aggregation, but migrates at an apparent molecular weight about 35% higher than the calculated. These observations suggest that SDS may bind anomalously to the antigenic core and that the isolated antigenic core may exhibit a propensity to aggregate under some solution conditions.

The preparation of R32LeuArg, R32tet₃₂, and R32NS1₈₁ has been described by others¹³; various degradation products have been isolated and identified. The present work addresses the development of assay methods for this family of antigens, focussing on the R32NS1₈₁ construct, and presents a discussion of some issues of general relevance to proteins derived from oligomeric repeating sequences. During the course of assaying samples for purity and stability, one probable degradative mechanism contributing to sample lability was identified, providing information useful to process optimization. No single method of monitoring purity and stability proved to be sufficient, but combined analysis by SDS-PAGE, Western blot analysis, and RPLC permitted the thorough characterization of samples of R32NS1₈₁.

EXPERIMENTAL

Materials

Trifluoroacetic acid (TFA), disodium ethylenediamine tetraacetic acid (EDTA), and 98–99% bovine serum albumin (BSA) were purchased from Sigma (St.

Louis, MO, U.S.A.). HPLC-grade water was prepared with a MilliQ water purification system (Millipore, Bedford, MA, U.S.A.). HPLC-grade acetonitrile, methanol, 30% hydrogen peroxide and 25% glutaraldehyde were from J. T. Baker (Phillipsburg, NJ, U.S.A.). Electrophoresis grade acrylamide, Coomassie Brilliant Blue R250 dye (CBB) and N,N'-methylene-bis-acrylamide (Bis) were purchased from Schwarz/Mann Biotech (Cleveland, OH, U.S.A.). N,N,N',N'-tetramethylethylenediamine (TEMED) and Bromphenol Blue were obtained from LKB (Paramus, NJ, U.S.A.). Sodium dodecyl sulfate (SDS) and 2-mercaptoethanol (BME) were from Pierce (Rockford, IL, U.S.A.). The silver stain kit was from ICN Radiochemicals (Irvine, CA, U.S.A.). The molecular weight markers [unstained No. 6000 LA or pre-stained No. 6041 LA, composed of a mixture of standard proteins of molecular weights 14.3, 18.4, 25.7, 43.0, 68.0, 97.4 and 200 kilodalton (kD)], ultrapure glycine, and trishydroxymethylaminomethane (Tris) were from Bethesda Research Labs. (Gaithersburg, MD, U.S.A.). Triton X-100 (electrophoretic grade), Tween 20, Protein A-horseradish peroxidase (HRP) conjugate and HRP color development reagent were from Bio-Rad Labs. (Rockville Centre, NY, U.S.A.). A pool of five monoclonal antibodies to R32tet₃₂ was prepared by Walter Reed Army Institute of Research (WRAIR) antibodies developed against NS1₈₁ were also obtained from WRAIR.

Methods

Chromatography. A Shimadzu ternary gradient system was used, consisting of LC-6A pumps, SIL-6A autoinjector, SCL-6A system controller, SPD-6A detector, and C-R3A integrator (Shimadzu, Kyoto, Japan). Column: Brownlee Aquapore butyl bonded stationary phase; 7 μm particle size silica in a 30 \times 4.6 mm I.D. cartridge (Brownlee Labs., Santa Clara, CA, U.S.A.). Chromatographic conditions: mobile phase A was 0.1% TFA in water, mobile phase B was 0.1% TFA in acetonitrile. All solvents were filtered through 0.45- μm filters (Rainin Instrument Co., Woburn, MA, U.S.A.). A linear gradient from 20 to 50% B over 25 min was used for the analytical and micropreparative work. Elution was performed at ambient temperatures. Sample preparation for HPLC: the protein stock solution, the concentration of which had been determined by amino acid analysis, was diluted with water to a concentration of 0.4 mg/ml. Effect of UV light was studied under controlled conditions (30°C, 500–2500 $\mu\text{W}/\text{cm}^2$) in a UV Chamber (Hotpack, Philadelphia, PA, U.S.A.).

Electrophoresis. The vertical electrophoresis system, including lower and upper buffer reservoirs, 16 cm \times 1.5 mm spacers, 16 \times 18 cm glass plates, a 20-toothed well-former, a No. 2297 Macrodrive 5 power supply, and a No. 2219 Multitemp II circulating bath were from LKB. The No. 2202 Ultrosan densitometer was from the same source. Electrophoretic conditions: SDS-PAGE electrophoresis, on 14% acrylamide [prepared from acrylamide-Bis (30:0.8)] was performed according to Laemmli¹². Samples were prepared under reducing conditions (5% BME), and either boiled for 3 min or incubated at room temperature as indicated. For gels used for Western blotting, prestained electrophoretic markers were used.

Silver staining¹⁴. The gel was treated with 2.5% glutaraldehyde, then washed three times with 400 ml water prior to staining with the silver stain kit.

Western blotting. Electrophoretic transfer onto 20 \times 20 cm nitrocellulose sheets (Bio-Rad Labs.) was done with a TE52 Transphor system (Hoefer Scientific Instruments, San Francisco, CA, U.S.A.). The Western blotting procedure was per-

formed according to refs. 15 and 16. Following transfer, the nitrocellulose was blocked with 3% BSA in a buffer of 0.2 M Tris, 8.8% sodium chloride and 0.4% EDTA solution, pH 7.5. The paper was washed with the same buffer containing 0.1% SDS and 0.05% Triton X-100 (but not BSA). The antibody was incubated in the Tris-EDTA-sodium chloride buffer to which had been added 0.05% Tween-20, then washed with the buffer alone. The paper was incubated with the Protein A-HRP conjugate in buffer, and the color was developed according to the manufacturer's directions.

RESULTS AND DISCUSSION

RPLC analysis of an experimental lot of R32NS1₈₁ resulted in a single major peak; four minor peaks are observed. Edman sequence analysis indicated that the N-terminus was essentially intact, but a small proportion of the antigen (initially estimated as 3%) was found to have a sequence beginning with Pro-Asn-Ala-Asn. The sample was also assayed by SDS-PAGE in a concentration series from 50 µg to 20 µg. On Coomassie staining, the preparation appeared pure by visual means, as shown in Fig. 1A. The major band migrated at an apparent molecular weight of 31 kD, and minor bands appeared at about 14, 24, 44 and 48 kD. Densitometry indicated the presence of other components at apparent molecular weights less than the principal band, but these were so faint as to be undetectable by visual means. On overlaying the Coomassie gel with silver stain, as is shown in Fig. 1B, the contrast was markedly enhanced, permitting the clear visualization of a manifold of bands at molecular weights less than the major band including a clear band at 19 kD. Two

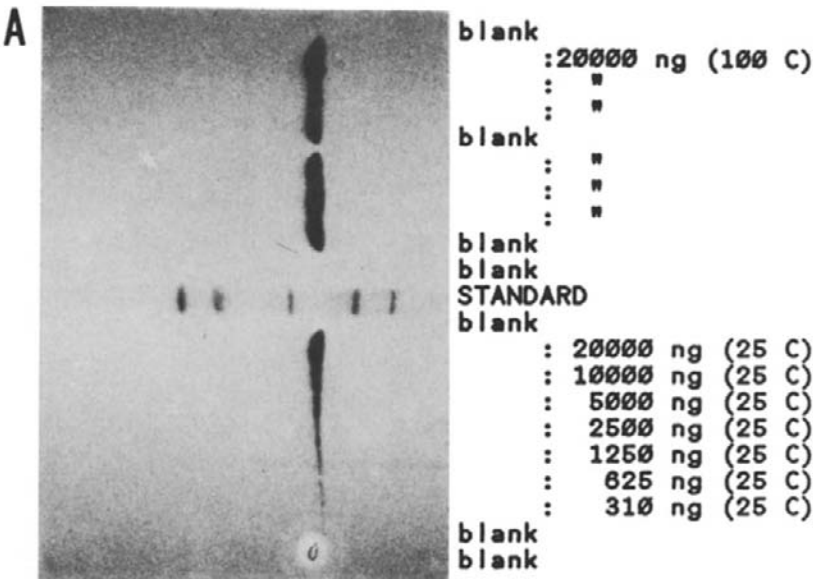


Fig. 1.

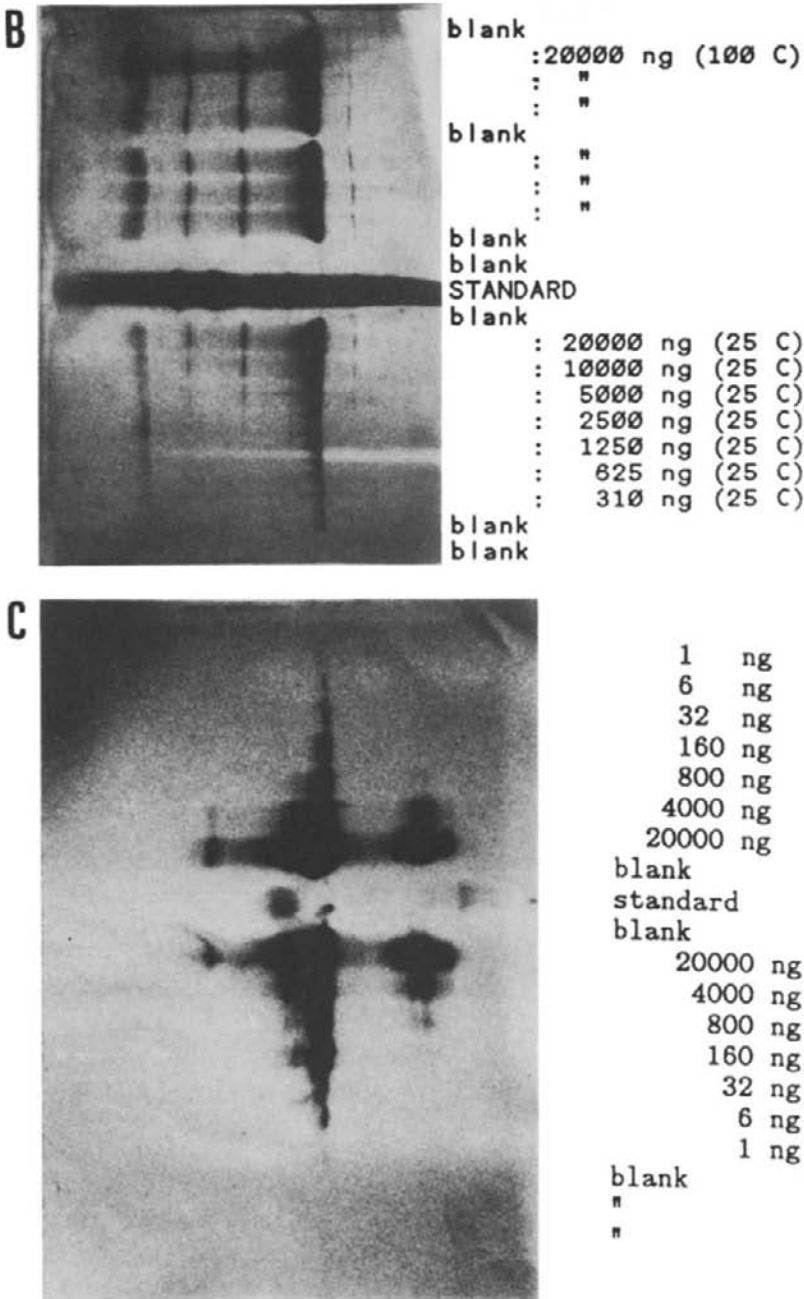


Fig. 1. SDS-PAGE of the R32NS1₈₁ malaria antigen: (A) gel after Coomassie Brilliant Blue (CBB) staining and (B) the result of a silver overlay. A major band can be detected at approximately 31 kD and minor bands at 14, 19, 24, 44 and 48 kD apparent molecular weight. A manifold of bands is observed in the region of 14–31 kD. (C) Western blot of an SDS-PAGE using anti R32tet₃₂ antibodies. The bands in the lane of the molecular weight standard represent 14, 19, 24, 31 and 48 kD. The labels 25C and 100C indicate that the disulfide reduction for electrophoresis was performed at 25°C or 100°C, respectively. Where not indicated, reduction was at 100°C.

other bands, at molecular weights of 44 and 48 kD were also visualized by silver staining. Western blotting with antibodies separately raised to the R32 and to the NS1₈₁ regions confirmed that the manifold of bands was antigen-derived. Interestingly, the R32 antibodies (see Fig. 1C) did not visualize the species at 14 kD, but NS1₈₁ antibodies did (data not shown). R32 antibodies also visualized a complex of bands at 54, 61, 76 and 97 kD that had not been visualized by any other means. This cluster of bands failed to react with NS1₈₁ antibodies, and was therefore believed to contain only R32. The failure to detect these protein bands by either Coomassie or silver staining is presumed to be due to the lack of basic or hydrophobic residues (or cysteine) in the R32 region. Two bands that were visualized by silver staining were not visualized by Western blotting, namely bands at 44 and 48 kD.

In an effort to investigate whether the apparent fragments of R32NS1₈₁ were real or whether they may have been generated during preparation for SDS-PAGE by boiling the protein samples in the presence of SDS sample buffer, SDS-PAGE of the RPLC fractions was performed.

Approximately 300 μ g of the antigen was fractionated by RPLC in 100 μ g portions. Fig. 2 shows a chromatogram of the region near the major peak, which eluted at 17.5 min. Four minor UV-positive peaks were resolved, representing a combined total of about 6% of the total peak area. Fourteen 0.5-ml fractions were collected, lyophilized, subjected to SDS-PAGE, and then visualized with silver stain as shown in Fig. 3. A number of those bands, notably those at apparent molecular weights of about 44 and 48 kD and those in the manifold of minor bands in PAGE were heterogeneously distributed through the different RPLC fractions, and are therefore not believed to be generated by the SDS-PAGE procedure. In one of the fractions, the manifold of minor bands was especially well resolved, and 24 resolved members of this ladder of bands were found to be at apparent molecular weight differences of about 580 D. Aside from the principal band, four bands were observed in each of the fractions: those at *ca.* 14, 15, 19 and 77 kD. Conceivably, these could have been generated by either exposure to the RPLC solvent or by sample prepa-

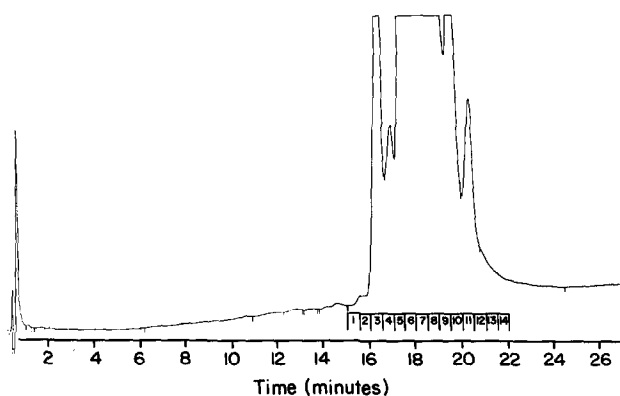


Fig. 2. Semi preparative RPLC separation of malaria antigen R32NS1₈₁ on a C₄ column, using a TFA-acetonitrile linear gradient. A 100- μ g aliquot of antigen was injected into the column and 14 fractions were collected for lyophilization and SDS-PAGE. Separations were performed at ambient temperature and elution was detected at 220 nm.

ration for SDS-PAGE, or perhaps they are simply not fractionated. The enormous complexity of the gel in Fig. 3 makes rigorous interpretation difficult, but certain conclusions can be reached. Apparently, RPLC has succeeded in partial fractionation of most of the manifold of components as seen by comparison of fractions 3–5 with 10–12 and 6–7 in which the principal band dominates. From fraction 6 to fraction 9, the average apparent molecular weights of the manifold of minor bands appears to decrease. The apparent molecular weight differences between adjacent bands of fraction 9 are about 580 D, not very different from the molecular weight of the R32 repeating unit (396 D for NANP and 425 D for NVDP). Therefore, it seems likely that the manifold of bands arose principally from cleavages in the R32 region, and that fragments of lower molecular weight were eluted later in RPLC. Since R32 is hydrophilic relative to NS1₈₁, the elution times are consistent with that hypothesis^{17,18}.

The analysis above suggested to us that the manifold of bands is probably principally derived from an Asp/Asn-Pro fragmentation pathway¹⁹. To test that hypothesis, samples were incubated under acidic conditions (which might be expected to catalyze Asn deamidation and Asp-Pro cleavage) and, separately, under a UV

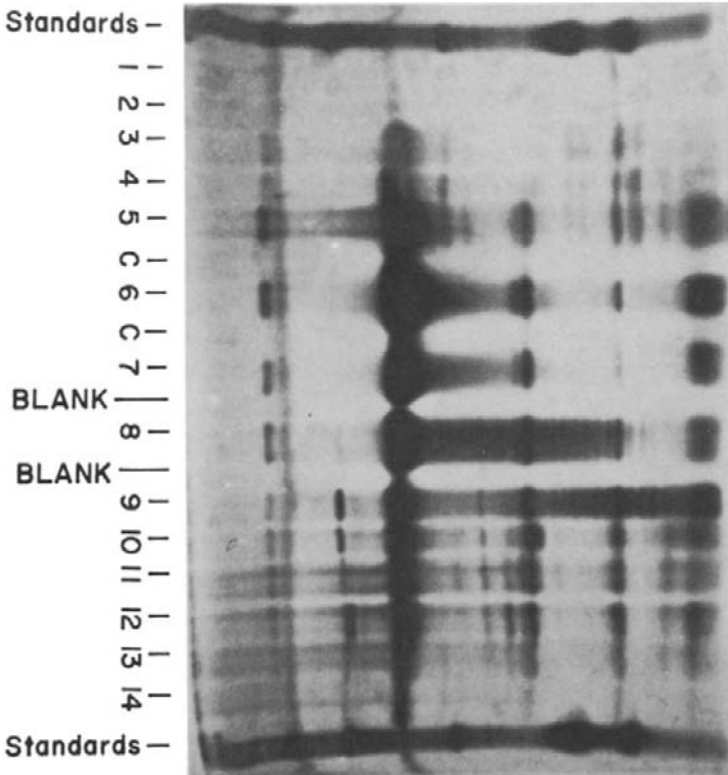


Fig. 3. SDS-PAGE of the collected fractions of the RPLC peak of R32NS1₈₁. The bands were visualized by silver staining. Comparison of fractions 3–5 with fractions 10–12 reveals obvious differences in the impurity distribution of the fractions.

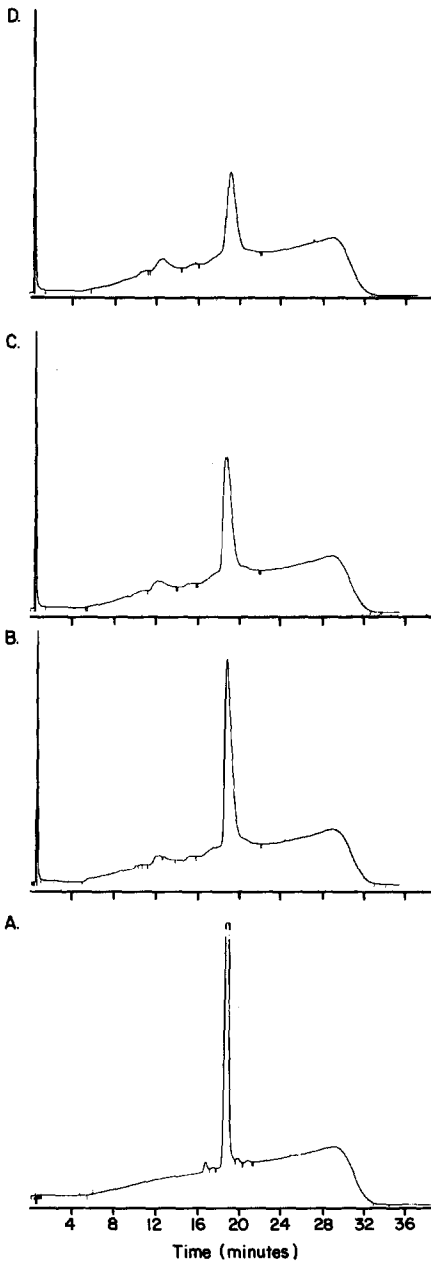


Fig. 4. The effect of acid treatment on the RPLC of R32NS1₈₁. A is the chromatogram of the control; B, C and D are the chromatograms of the antigen after 24, 72 and 96 h of incubation in 1 M hydrochloric acid. Chromatographic conditions are as in Fig. 2, except that 10 μ g antigen was injected into the column.

source (which might be expected to accelerate degradation in the tryptophan-containing NS1₈₁ region). Also, samples were incubated at 85°C to study the effect of simple thermal degradation. Fig. 4A shows a chromatogram of an untreated sample

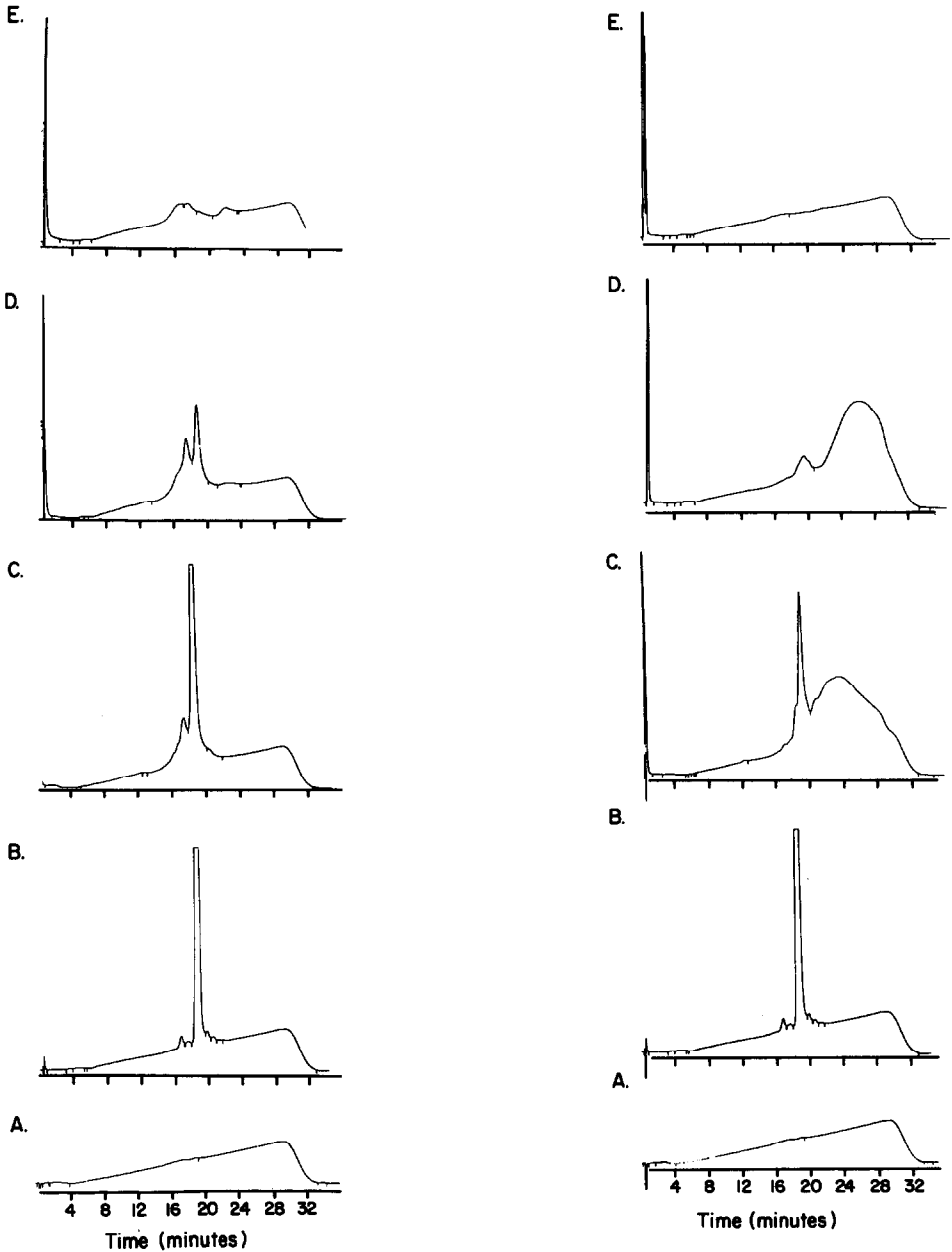


Fig. 5. Effect of UV light on the RPLC of R32NS1₈₁. A and B represent the chromatograms of a blank gradient and a control respectively, while C, D and E are the chromatograms of samples exposed to UV light for 6 h, 1 day and 5 days respectively. Chromatographic conditions are as in Fig. 4, except that 10 μ g antigen was injected into the column.

Fig. 6. Effect of incubation at 85°C on the RPLC of R32NS1₈₁₂. A and B represent the chromatograms of a blank gradient and a control respectively, while C, D and E are the chromatograms of samples exposed to 85°C for 6 h, 1 day and 5 days respectively. Chromatographic conditions are as in Fig. 5, except that 10 μ g antigen was injected into the column.

as a control; B, C and D are chromatograms of the antigen after 24, 72 and 96 h of incubation with 1 *M* hydrochloric acid at room temperature. It is apparent that the peakwidth increases with increasing exposure to acidic conditions. The appearance of a peak at t_0 and a broad peak at about 12 min (which corresponds to the elution position of R32LeuArg) is also observed in acid-catalyzed degradation.

The effect of UV light on the RPLC behavior of R32NS1₈₁ is shown in Fig. 5. The antigen solution was stored in UV-transparent tubes in a standard chamber, and aliquots were analyzed by RPLC and SDS-PAGE. Fig. 5A and B represents the chromatograms of a blank gradient and a control, while C, D and E are chromatograms after exposure to UV for 6 h, 1 day and 5 days, respectively. Due to UV degradation, the sharp RPLC peak is split into two closely spaced peaks with similar retention times. Eventually, the antigen peak is abolished, and a t_0 peak appears. The corresponding SDS-PAGE analysis shows an increased band intensity at apparent molecular weights both greater than and less than that of the antigen.

Finally, the effect of thermal degradation was explored. Samples were incubated at 85°C. Fig. 6A and B shows a blank gradient and a control, respectively, while C, D and E are chromatograms of samples treated for 6 h, 1 day and 5 days, respectively. A broad peak appears at a retention time greater than that of the antigen. Eventually, only a t_0 peak is observed. On SDS-PAGE, it appears that the products of heat treatment represent a low-molecular-weight distribution of species.

To summarize the results of this work, numerous bands observed on SDS-PAGE proved to be possible to fractionate on RPLC and were therefore not artifacts of sample preparation. Because the spacing of the lower-molecular-weight manifold bands was consistent with the molecular weight of the tetrapeptide repeat, these were suggested to be derived from the sequential cleavage of the R32 portion of the antigen. Sequence analysis, which may have underestimated the amount of degradation, was consistent with Asp-Pro cleavage. Degradation studies under acidic conditions, exposure to UV light or an incubation at elevated temperatures gave results that support of this hypothesis.

CONCLUSIONS

The apparently simple RPLC elution profile represents a complex mixture of impurities as seen by SDS-PAGE. The selectivity of RPLC is insufficient to separate this mixture into its individual components. RPLC does resolve the 44- and 48-kD components (see fractions 9–12 in Fig. 3) from the main peak. Highly purified R32NS1₈₁ for clinical trials could be obtained by taking a center cut of the RPLC peak. Degradation of the protein resulted in significant RPLC peak broadening without the concomitant appearance of specific degradate peaks. Judicious consideration of data from multiple sources can suggest a mechanism by which degradation products arise in a sample. In the present work, the data were consistent with an acid-catalyzed deamidation/Asp-Pro cleavage mechanism, generating a manifold of bands separated in apparent molecular weight by about 580 D. Both NS1₈₁ and R32 were apparently generated by the process. Purity determinations by SDS-PAGE should be approached with caution, since differences in loading and sample preparation can cause assay variability. Differentiating artifacts from impurities requires patience and well-controlled experiments.

Apparent molecular weights determined by SDS-PAGE may be utterly anomalous and misleading. In the present work, the R32 core apparently binds SDS poorly and/or aggregates, and therefore migrates at a much higher apparent molecular weight than R32NS1₈₁. In general, one would, therefore, expect that cleavage within the NS1₈₁ region would generate a fragment of higher apparent molecular weight than the parent, as well as a fragment small enough to migrate at the dye front. On the other hand, cleavage within the R32 region would generate a stainable fragment of lower apparent molecular weight, and an unstainable fragment of high-molecular-weight.

No single technique is adequate to characterize this sample fully. For the present work, both CBB and the silver staining method used here failed to detect R32, which was detected by Western blotting. On the other hand, the antibodies used in these Western blotting experiments did not detect trace impurities (at 44 and 48 kD) which were detected by conventional electrophoretic stains. The verdict for the performance of RPLC is ambiguous. It did give a reasonable estimate of the purity of the sample, but the existence of peaks appears to have been essentially by chance, and not indicative of specific impurities.

The above described careful and systematic analytical approach has been routinely applied to the characterization of the different DNA derived antigens produced for clinical trials.

ACKNOWLEDGEMENTS

The authors wish to thank J. Young and M. Gross for the cloning and expression of the R32 constructs; W. R. Ballou and M. Scott for the monoclonal antibodies; D. Zabriskie and D. Wareheim for fermentations; R. Sitrin, J. Dingerdissen and G. Folea Wasserman for the purification processes; and I-Yih Huang for the amino acid and N-terminal analysis. Special thanks are due to T. R. Malefy, W. Vencek, E. R. White and J. L'Italien for helpful discussions in the preparation of the manuscript.

REFERENCES

- 1 R. S. Nussenzweig, J. Vanderberg and H. Most, *Mil. Med.*, 134 (1969) 1176.
- 2 D. F. Clyde, V. McCarthy, R. M. Miller and W. E. Woodward, *J. Trop. Med. Hyg.*, 24 (1975) 397.
- 3 L. H. Miller, R. J. Howard, R. Carter, M. F. Good, V. Nussenzweig and R. Nussenzweig, *Science (Washington, D.C.)*, 234 (1986) 1349.
- 4 D. Mazier, S. Mellouk, R. L. Beaudoin, B. Texier, P. Druilhe, W. Hockmeyer, J. Trosper, C. Paul, Y. Charoenvit, J. Young, F. Miltgen, L. Chedid, J. P. Chigot, B. Galley, O. Brandicourt and M. Gentilini, *Science (Washington, D.C.)*, 231 (1986) 156.
- 5 N. Yoshida, R. S. Nussenzweig, P. Potocnjak, V. Nussenzweig and M. Aikawa, *Science (Washington, D.C.)*, 207 (1980) 71.
- 6 F. Zavala, J. P. Tam, M. R. Hollingdale, A. H. Cochrane, I. Quakyi, R. S. Nussenzweig and V. Nussenzweig, *Science (Washington, D.C.)*, 228 (1985) 1436.
- 7 M. R. Hollingdale, E. H. Nardin, S. Tharavanij, A. L. Schwartz and R. Nussenzweig, *J. Immunol.*, 132 (1984) 909.

- 8 J. B. Dame, J. L. Williams, T. F. McCutchan, J. L. Weber, R. A. Wirtz, W. T. Hockmeyer, W. L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, G. S. Sanders, E. P. Reddy, C. L. Diggs and L. H. Miller, *Science (Washington, D.C.)*, 225 (1984) 593.
- 9 J. F. Young, W. T. Hockmeyer, M. Gross, W. R. Ballou, R. A. Wirtz, J. H. Trosper, R. L. Beaudoin, M. R. Hollingdale, L. H. Miller, C. L. Diggs and M. Rosenberg, *Nature (London)*, 228 (1985) 958.
- 10 W. R. Ballou, J. A. Sherwood, F. A. Neva, D. M. Gordon, R. A. Wirtz, G. F. Wasserman, C. L. Diggs, S. L. Hoffman, M. R. Hollingdale, W. T. Hockmeyer, I. Schneider, J. F. Young, P. Reeve and J. D. Chulay, *Lancet*, ii (1987) 1277.
- 11 J. K. Swadesh, unpublished results.
- 12 U. K. Laemmli, *Nature (London)*, 227 (1970) 680.
- 13 G. F. Wasserman, R. Inacker, C. Cohen Silverman and M. Rosenberg, in R. Burges (Editor), *UCLA Symposia on Molecular and Cellular Biology, Protein Purification Micro to Macro*, Vol. 68, Allen R. Liss, New York, 1987, p. 337.
- 14 B. R. Oakley, D. R. Kirsch and N. R. Morris, *Anal. Biochem.*, 105 (1980) 361.
- 15 H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 4350.
- 16 J. Renart, J. Reiser and G. R. Stark, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 3116.
- 17 J. M. R. Parker, D. Guo and R. S. Hodges, *Biochemistry*, 25 (1986) 5425.
- 18 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499.
- 19 M. Landon, *Methods Enzymol.*, 47 (1977) 145.