

ON THE LOCATION AND STRUCTURE OF THE ACTIVE SITES OF ANTIBODY MOLECULES*

BY S. J. SINGER AND NEAL O. THORPE†

DEPARTMENT OF BIOLOGY, UNIVERSITY OF CALIFORNIA, SAN DIEGO (LA JOLLA)

Communicated by Linus Pauling, May 17, 1968

By the method of affinity-labeling¹ it has been possible to attach chemical tags to amino acid residues that, by a number of stringent criteria, are contact residues within the active sites of antibody (Ab) molecules.^{2, 3} With IgG Ab specific to three different benzenoid haptens from four different mammalian species,^{3, 4} it was found that tyrosine residues on both heavy (H) and light (L) chains were affinity-labeled by the specific diazonium reagents used. These results were interpreted³ to mean that structurally unique and characteristic tyrosine residues were present in all the active sites that were affinity-labeled. To identify these residues, we have recently isolated and characterized dipeptide fragments bearing the affinity label from the H and L chains of rabbit and mouse IgG Ab specific for the 2,4-dinitrophenyl (DNP) group. Our results (1) demonstrate that the labeled residues are indeed unique; (2) establish that the variable segment of the L chain⁵ is directly involved in forming the Ab active site; (3) strongly suggest that the labeled tyrosine is residue 86 from the amino-terminal end of the L chain; and (4) lead to the suggestion that in the Fab fragment (and, in particular, in the active site) of each half of an IgG molecule, the H and L chains are structurally related by a dyad axis of pseudosymmetry.⁶

Materials and Methods.—Rabbit anti-DNP Ab, elicited by immunization with DNP-bovine γ -globulin, were purified⁷ and affinity-labeled with H³-*m*-nitrobenzenediazonium fluoroborate (H³-MNBDF),⁸ and the labeled H and L chains were isolated,⁹ as in our previous studies. Mouse anti-DNP Ab were raised to DNP-keyhole limpet hemocyanin in Swiss-Webster mice in which Ehrlich ascites tumor cells were intraperitoneally implanted.¹⁰ The mouse anti-DNP Ab were isolated from the ascitic fluid and affinity-labeled with H³-MNBDF by methods, and with results, analogous to those for rabbit Ab.

The labeled chains were digested with Nagarse enzyme (Enzyme Development Corp.) at a weight ratio of enzyme to peptide of 1:50 in 0.2 M NH₄HCO₃ for 20 hr at 40°C. Gel filtration of the digests was carried out on P-2 Biogel in 0.05 M NH₄HCO₃, and the elution of radioactivity was monitored. The elution profiles obtained with rabbit-chain digests have been presented elsewhere.⁶ The profiles for the mouse-chain digests are shown in Figure 1. In all cases, a major fraction of radioactivity, accounting for about 50% of the original affinity label, was obtained. This major fraction, according to gel filtration experiments with calibrated P-2 Biogel columns in phenol:acetic acid:water (1:1:1), was predominantly dipeptide in size (see also below). Very little of the radioactivity was released as free *m*-nitrobenzeneazotyrosine.¹¹ The radioactive peptides from this major fraction were isolated in a pure state and in high yield by taking advantage of the fact that the *m*-nitrobenzeneazotyrosine-labeled peptides are haptens capable of specific reversible binding to the active sites of anti-DNP Ab. The labeled peptide fraction was mixed with unlabeled rabbit anti-DNP Ab in the molar ratio of 1:1, and the mixture was brought to half-saturation in (NH₄)₂SO₄. The precipitate that formed was washed and then dissolved in dilute buffer; a slight excess of DNP-aminocaproate was added to release the labeled peptides, and the Ab was reprecipitated with (NH₄)₂SO₄. The supernatant solution contained about 90% of the radioactivity of the original peptide fraction. The details of the purification procedure used will be given elsewhere.¹²

In the final step of the procedure, the Ab-isolated peptides were chromatographed on Beckman amino acid analyzer resin PA-35 and eluted by a pyridine-acetate gradient.¹³

The results for the mouse Ab chains are shown in Figure 2*a* and *b*; those for the rabbit-chain peptides have already been presented.⁶ The fractions indicated were pooled and subjected to amino acid analyses. The results of these analyses, as well as corresponding ones for the labeled dipeptides from rabbit H and L chains, are given in Table 1. For the rabbit chains, which were available in sufficient amounts to validate these procedures, it was shown that the amino acid analyses in Table 1 were representative of independent preparations of the labeled peptides.

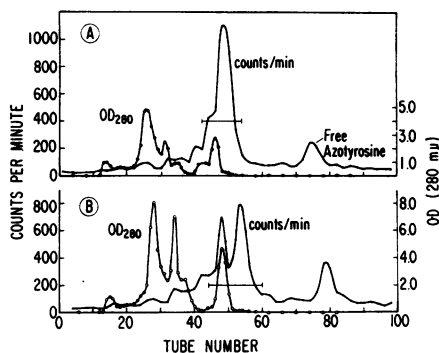


FIG. 1.—Nagarse digests of (A) L chains and (B) H chains of affinity-labeled mouse anti-DNP Ab separated by gel filtration on a column (95 × 2.5 cm) of P-2 Biogel equilibrated with 0.05 M NH_4HCO_3 .

The elution profile of radioactivity (filled circles) marks the distribution of affinity label. The free azotyrosine was identified with the aid of an authentic sample of the compound H^3 -*m*-nitrobenzeneazotyrosine. The fractions indicated by the bars were pooled and the labeled peptides were subsequently isolated from them by the Ab purification method (see text).

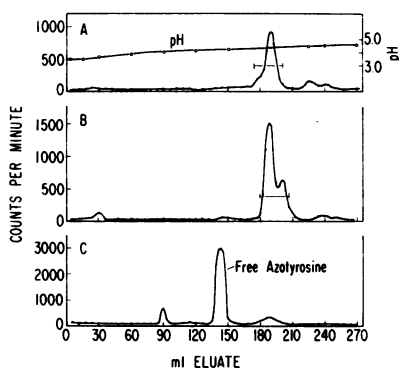


FIG. 2.—Chromatographic separations on a column (12 × 0.9 cm) of Beckman custom research resin type PA-35 maintained at 50°C and eluted with a pyridine-acetate pH gradient at a pumped rate of 60 ml/hr. The pH gradient is shown only in (A) but was the same for (B) and (C).

(A) and (B) show the elution profiles of radioactivity for the purified affinity-labeled dipeptides from the H and L chains, respectively, isolated from the fractions in Fig. 1. The fractions indicated by the bars were pooled for the amino acid analyses of Table 1.

(C) shows the elution of radioactivity after the purified peptides from the L chains (B) were subjected to one step of an Edman degradation. After the peptides were coupled with excess phenylisothiocyanate and the products subjected to the cyclization and cleavage reactions, the entire mixture was passed through the column and the radioactivity of the affinity label was monitored. The position of the free azotyrosine peak was determined with an authentic sample of the compound H^3 -*m*-nitrobenzeneazotyrosine. The minor radioactive peak at 90-ml elution volume is H^3 -*m*-nitroaniline, which results from a small amount of cleavage of the azo-linkage.

It was demonstrated by Edman degradation experiments that the purified labeled peptides in all cases were dipeptides. After the labeled peptides were subjected to a one-step Edman cleavage, over 90% of the radioactivity of the peptides was identified as free H^3 -*m*-nitrobenzeneazotyrosine (Fig. 2c). The labeled tyrosine residue must therefore have been the carboxyl-terminal residue of the labeled dipeptides.

Results and Discussion.—From the amino acid analyses in Table 1, and the results of the Edman degradation studies, we conclude that from each chain of affinity-labeled anti-DNP Ab, a set of labeled dipeptides can be recovered that accounts for a large fraction of the original label on the chain. Within each set, one labeled dipeptide generally predominates:¹⁴ it is -aspartyl- (or asparaginyll-) tyrosyl- for mouse L chains; -valyl-tyrosyl- for rabbit L chains; and -threonyl-tyrosyl- for rabbit H chains. These major dipeptides account for at least 50 per cent of the label within each set. In addition, smaller amounts of several other dipeptides are found within each set. The possible significance of these minor dipeptides is discussed elsewhere.¹² For our present purposes, the presence of a single dominant dipeptide (a different one in each case) strongly suggests that a unique and characteristic tyrosine residue is affinity-labeled on most, if not all, of the molecules present in each chain preparation.

To identify these unique tyrosine residues structurally, the most useful starting point is a comparison of the dipeptide data obtained for mouse Ab L chains with amino acid sequence data for mouse α -type Bence-Jones proteins, two of which have been almost completely sequenced.¹⁵ The studies of Hilschmann and Craig⁵ first revealed the fact that L chains consist of two distinguishable but connected segments: a carboxyl-terminal half which is essentially invariant for all chains of a given species and class, and an amino-terminal half which is variable in composition and sequence from one L chain to the next. There are four tyrosine residues in the invariant segments of the two mouse Bence-Jones proteins studied. None of these has an aspartyl or asparaginyll residue as its amino-terminal neighbor. The affinity-labeled tyrosine on mouse L chains therefore cannot be in the invariant segment of the chains and must be in the variable segment. Since the evidence is very strong^{2, 3} that the labeled tyrosine is a contact residue in the Ab active site, these results demonstrate directly that the Ab active site is contained, at least partly, within the variable segment of the L chain. This has been inferred ever since the existence of the variable segment was revealed, but it has not been as directly demonstrated heretofore.

A definitive localization of the labeled tyrosine residue within the variable segment of the L chain is difficult at present because only the two mouse Bence-Jones sequences are known.¹⁵ An identification can be made provisionally, however, which is consistent with the facts available and which has interesting corollaries for the structure of Ab molecules. These two mouse Bence-Jones proteins (designated MBJ-41 and MBJ-70) have in their variable segments the tyrosine residues listed in Table 2. Tyrosine-49 has isoleucine as its amino-terminal neighbor in both chains, as well as in all human α and λ Bence-Jones proteins so far examined;¹⁶⁻²⁰ it is therefore not likely to be the labeled residue in mouse Ab L chains. Tyrosine-91, present in MBJ-41 but not in MBJ-70, has glutamine as its amino-terminal neighbor; glutamine is also present at position 90 in MBJ-70 and probably in all but one of the human α Bence-Jones proteins so far examined. Tyrosine-91 is therefore also an unlikely candidate. Tyrosine-

TABLE 1. *Amino acid analyses of purified labeled dipeptides from affinity-labeled antibody chains.**

Amino acid†	μMoles Recovered			
	Mouse L	Mouse H	Rabbit L‡	Rabbit H‡
Asp	0.052	0.008	Trace	0.005
Thr	<i>0.015</i>	<i>0.011</i>	Trace	0.050
Ser	<i>0.011</i>	<i>0.019</i>	0	<i>0.013</i>
Glu	0.006	0.005	0.005	0.002
Gly	0.005	<i>0.016</i>	0.003	<i>0.013</i>
Ala	0.002	0.005	0.002	0.005
Val	0.002	0.003	0.041	0.003
Label applied (μM)§	0.090	0.054	0.078	0.098
Amino acids recovered (μM)§	0.101	0.078	0.055	0.097

* Hydrolysis was performed with a measured amount of labeled peptide (as determined by H³ content) in 6 N HCl at 110° for 20 hr. The values recorded are corrected for the small amounts of contaminating free amino acids in the unhydrolyzed samples, but are not corrected for destruction during hydrolysis.

† For brevity, those amino acids that did not exceed 0.003 μmole in any of the four analyses are not included in the table. The values in *boldface* designate the predominant amino acid, the values in *italics* those amino acids present in lesser but significant amounts. The labeled tyrosine residue is destroyed on acid hydrolysis and no ninhydrin-positive product is detected, as shown by experiments with the model compound, H³-*m*-nitrobenzeneazotyrosine.

‡ Analysis shown is for the major peak appearing on elution of the labeled dipeptides from the PA-35 resin column (see Fig. 3 of ref. 6).

§ Because of the destruction of the labeled tyrosine residue on acid hydrolysis, the mixture of labeled dipeptides should have yielded 1 mole of total amino acids per mole of label. The values for the μmoles of amino acids recovered includes the trace amounts of those amino acids not included in the table.

71 in MBJ-41 and tyrosine-86 in the two proteins, however, are both possibilities, since aspartic acid residues may occur as their amino-terminal neighbors.

We do not have evidence at the present time that discriminates unambiguously between these two tyrosines. However, it is an interesting fact that tyrosine-86 appears in all 12 mouse κ and human κ and λ Bence-Jones proteins so far sequenced in that region (Table 2), whereas tyrosine-71 has so far occurred only in 1 out of 10. In other words, tyrosine-86 appears to be a residue that is constant within the variable segments of L chains. The assignment of the labeled tyrosine as tyrosine-86 would therefore be consistent with the inference that we arrived at earlier,³ namely, that it is the same L-chain tyrosine residue that is affinity-labeled in different Ab active sites, and that this tyrosine residue must be a constant feature of different Ab L chains.

Of further interest is the fact that for the 12 Bence-Jones proteins listed in Table 2, the 3 most commonly observed amino acids at position 85 are aspartic acid (or perhaps asparagine), valine, and threonine. It is remarkable that these are the amino acids that are the predominant amino-terminal neighbors to the labeled tyrosine residue in mouse L, rabbit L, and *rabbit H* chains, respectively, of anti-DNP Ab.

A hypothesis^{3, 6} that provides a consistent and unifying explanation for all of our affinity-labeling results is that: (1) most, if not all, of the L chains of the rabbit, mouse, and human species have an evolutionarily conserved tyrosine residue at position 86 (or its equivalent) from the amino-terminal ends of the chains; (2) H chains are not only evolutionarily related to L chains,^{3, 21, 22} but also have the homologous tyrosine residue at position 86 (or its equivalent); (3) both this tyrosine residue on an L

of chain, there is a striking similarity of these labeled peptides from Ab of the two specificities, which is consistent with the proposition that the same tyrosine residues are affinity-labeled independent of the specificity of the Ab.

This hypothesis suggests further the existence of certain symmetry properties⁶ in Ab active sites and in Fab fragments of IgG molecules, which are schematically represented in Figure 3. Features that are incorporated into this structure, and

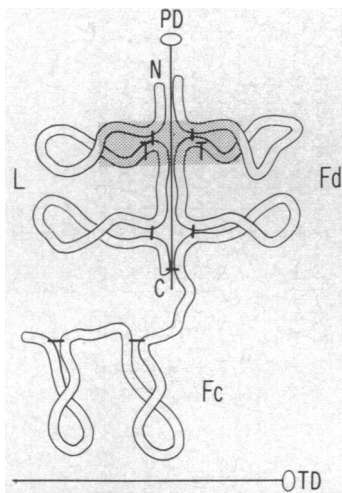


FIG. 3.—A representation of the proposed symmetry relationships in each half of an IgG immunoglobulin molecule. (*TD* represents the true dyad axis, which presumably relates the two identical halves of the whole molecule.) The L chain and Fd fragment of the H chain are related by a dyad axis of pseudosymmetry (*PD*). *N* and *C* refer to the amino- and carboxyl-terminals, respectively, of the L chain. The solid bars between chains and between chain segments indicate the disulfide bridges within the half molecule. The stippled region very schematically represents the Ab active site of the half molecule; it is intended to show the general location of the active site and illustrate the proposal that the *PD* axis relates the portions of the active site contributed by the L and H chains. The *T* symbols denote the tyrosine residues on the L and H chains, which, it is proposed, are the unique residues that become affinity-labeled in the active sites (see text for details).

that derive from studies other than those reported in this paper, include: (1) the structural homologies that probably exist among the different, approximately 100-amino-acid-long, variable and invariant segments of H and L chains;^{21, 22} (2) a folding of these segments that is required by their intrachain -S-S- bridges;²⁸ and (3) the fact that the cysteine residue on the H chain and the cysteine residue on the L chain, which form the single disulfide bridge between the chains, occur within homologous stretches of amino acid sequence²⁹ and are therefore presumably located at homologous positions from the amino-terminal ends of their respective chains. It is proposed that a twofold axis of pseudosymmetry relates the L chain and the nonidentical Fd portion of an H chain, which make up a Fab fragment, in a manner roughly analogous to the twofold pseudoaxis that relates the α and β chains of half of a hemoglobin molecule.³⁰ (This is in addition to a true dyad axis that presumably relates the two identical halves of an IgG molecule). The twofold pseudoaxis bisects and is perpendicular to the -S-S-bridge connecting the H and L chains and passes through the Ab active site so that those portions of the site that are contributed by the L chain and the Fd piece are not only homologous to one another in sequence but are symmetrically disposed about the pseudoaxis. Finally, it is proposed that the homologous tyrosine residues at positions 86 on the H and L chains that are affinity-labeled are not only proximal to each other but are pseudosymmetrically related. This last feature provides a structural explanation for the fact that the ratio of affinity label found on H and L chains is nearly constant for different Ab,³ for it suggests that the affinity-labeling reagent, when reversibly bound to a site, may occupy a position along the axis of pseudosymmetry such that there is a relatively constant probability for it to react with either the H- or the L- chain tyrosine in the site.

The proposed pseudosymmetry of the Fab fragment is related to the pairing of H and L chains to form an IgG molecule. The discovery of the existence of H and L chains in each Ab molecule originally led Edelman and Benacerraf³¹ to suggest that both chains might be involved in the determination of the specificity of the Ab. An attractive feature of this proposal was the possibility that within a single cell, H and L chains might arise independently of one another and then pair to give a molecule of a particular Ab specificity. By this means, the diversity of Ab specificities would be greatly amplified over that which could be achieved by one chain alone.

At an earlier stage in our affinity-labeling studies, however, it was suggested³ that the H and L chains produced by a single cell might not arise independently. The argument that was employed may be stated in terms of the more detailed structure shown in Figure 3. The conformations of the variable segments of the H and L chains making up an Ab molecule might have to be correlated in order to achieve the required symmetry relationship between them in the intact molecule. For example, the proper pseudosymmetry in the active site might not result unless it were permitted by the conformations adopted by the *individual* H and L chains.

Other investigations have indicated that the variable segments of the H and L chains of an Ab molecule are structurally correlated. Chain hybridization experiments,³² for example, show quite convincingly that H and L chains from rabbit IgG Ab molecules of a given specificity bind to one another more firmly than do chains from different molecules. (Similar results had earlier been obtained with H and L chains from human myeloma immunoglobulins.³³) The *preferential* binding of homologous H and L chains to one another is to be attributed to interactions between the variable segments of these chains.³² A special structural relationship between the variable segments of H and L chains from the same immunoglobulin molecule is also suggested by the specific optical rotatory effects that accompany the recombination of the chains.³⁴

There is no *a priori* reason why the specificity of an Ab molecule and the binding affinity of its H and L chains should be correlated. Crudely put, the amino acid residues in an Ab active site are there to make contact with the antigenic determinant, not with each other; the residues primarily responsible for inter-chain bonding must be largely outside the active site (cf. Fig. 3).

If, therefore, the variable segments of H and L chains from a given IgG molecule are structurally correlated, it is likely that there exists some substantial relationship between the amino acid sequences of these variable segments. Such a relationship goes beyond a simple evolutionary one. This in turn implies that the generation of H and L chains in a given immunoglobulin-producing cell is not independent but is somehow coordinated. Either a direct coordination may exist between the somatic mutational processes that generate H- and L-chain multiplicity³ or, as suggested by Cohn,³⁵ some special selective process (independent of the antigen) may operate to stabilize those cells that generate H and L chains that have an especially large binding affinity for one another.

We gratefully acknowledge the invaluable assistance of Mrs. Birgitta Kiefer. We also benefited greatly from discussions with Dr. Russell F. Doolittle and Mr. John W. Fenton, II.

* This work was supported by grant AI-06659 from the National Institute of Allergy and Infectious Diseases, U.S. Public Health Service.

† Advanced research fellow of the American Heart Association, 1966–1967. Postdoctoral fellow of the National Institutes of Health, 1965–1966. Present address: Department of Biology, Augsburg College, Minneapolis, Minnesota 55404.

- ¹ Wofsy, L., H. Metzger, and S. J. Singer, *Biochemistry*, **1**, 1031 (1962).
- ² Metzger, H., L. Wofsy, and S. J. Singer, *Biochemistry*, **2**, 979 (1963).
- ³ Singer, S. J., and R. F. Doolittle, *Science*, **153**, 13 (1966).
- ⁴ Good, A. H., Z. Ovary, and S. J. Singer, *Biochemistry*, **7**, 1304 (1968).
- ⁵ Hilschmann, N., and L. C. Craig, these PROCEEDINGS, **53**, 1403 (1965).
- ⁶ Singer, S. J., L. I. Slobin, N. O. Thorpe, and J. W. Fenton, II, in *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 32 (1967), p. 99.
- ⁷ Farah, F. S., M. Kern, and H. N. Eisen, *J. Exptl. Med.*, **112**, 1195 (1960).
- ⁸ Good, A. H., P. S. Traylor, and S. J. Singer, *Biochemistry*, **6**, 873 (1967).
- ⁹ Fleischman, J. B., R. H. Pain, and R. R. Porter, *Arch. Biochem. Biophys.* (Suppl. 1), 174 (1962).
- ¹⁰ For introducing us to this technique and for the generous donation of the Ehrlich tumor cells, we are greatly indebted to Drs. Leon Wofsy and Benjamin Burr.
- ¹¹ Digestion of the labeled chains with pronase or with leucine aminopeptidase, however, released most of the label as *m*-nitrobenzeneazotyrosine.
- ¹² Thorpe, N. O., and S. J. Singer, in preparation.
- ¹³ Benson, J. V., Jr., R. T. Jones, J. Cormick, and J. A. Patterson, *Anal. Biochem.*, **16**, 91 (1966).
- ¹⁴ No one amino acid was clearly predominant in the peptides from the mouse Ab H chains. This may well have been due to the fact that appreciable amounts of the different classes $\gamma 1$, $\gamma 2a$, and $\gamma 2b$ of H chains were present, according to immunoelectrophoretic analyses. (Dr. L. Herzenberg kindly performed these analyses for us.) It is possible that the three principal labeled dipeptides found, -seryl-tyrosyl-, -glycyl-tyrosyl-, and -threonyl-tyrosyl-, corresponded to the three different H-chain classes. The other chain populations in Table 1 were most likely of a single predominant class.
- ¹⁵ Gray, W. R., W. J. Dreyer, and L. Hood, *Science*, **155**, 465 (1967).
- ¹⁶ Hilschmann, N., *Z. Physiol. Chem.*, **348**, 1077 (1967).
- ¹⁷ Putnam, F. W., K. Titani, and E. Whitley, Jr., *Proc. Roy. Soc. (London)*, Ser. B, **166**, 124 (1966).
- ¹⁸ Milstein, C., *Nature*, **216**, 330 (1967).
- ¹⁹ Dreyer, W. J., W. R. Gray, and L. Hood, in *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 32 (1967), p. 333.
- ²⁰ Putnam, F. W., T. Shinoda, K. Titani, and M. Wilber, *Science*, **157**, 1050 (1967).
- ²¹ Hill, R. L., R. Delaney, H. E. Lebovitz, and R. E. Fellows, Jr., *Proc. Roy. Soc. (London)*, Ser. B., **166**, 159 (1966).
- ²² Singer, S. J., and R. F. Doolittle, in *Structural Chemistry and Molecular Biology*, eds. A. Rich and N. Davidson (San Francisco: Freeman, 1968), p. 173.
- ²³ Frangione, B., and C. Milstein, *Nature*, **216**, 939 (1967).
- ²⁴ Pink, J. R. L., and C. Milstein, *Nature*, **216**, 941 (1967).
- ²⁵ An examination of this one H-chain sequence, and of the L-chain sequences in Table 2, leads to the prediction that if labeled *tripeptides* could be isolated from the H and L chains of affinity-labeled mouse and rabbit anti-DNP Ab that contained the labeled tyrosine-86 residue as their carboxyl-terminals, then their *N*-terminal residues might be predominantly alanine (position 84).
- ²⁶ Fenton, J. W., II, and S. J. Singer, *Biochem. Biophys. Res. Commun.*, **20**, 315 (1965).
- ²⁷ Fenton, J. W., II, Ph.D. thesis, University of California at San Diego, La Jolla (1968).
- ²⁸ Milstein, C., *Nature*, **209**, 370 (1966).
- ²⁹ Steiner, L. A., and R. R. Porter, *Biochemistry*, **6**, 3957 (1967).
- ³⁰ Muirhead, H., J. M. Cox, L. Mazzarella, and M. F. Perutz, *J. Mol. Biol.*, **28**, 117 (1967).
- ³¹ Edelman, G. M., and B. Benacerraf, these PROCEEDINGS, **48**, 1035 (1962).
- ³² Roholt, O. A., G. Radzinski, and D. Pressman, *J. Exptl. Med.*, **125**, 191 (1967).
- ³³ Grey, H. M., and M. Mannik, *J. Exptl. Med.*, **122**, 619 (1965).
- ³⁴ Dorrington, K. J., M. H. Zarlengo, and C. Tanford, these PROCEEDINGS, **58**, 996 (1967).
- ³⁵ Cohn, M., in *Nucleic Acids in Immunology*, ed. V. Bryson and H. J. Vogel (New York: Academic Press), in press.