Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 260, No. 11, Issue of June 10, pp. 6514-6517, 1985

The M_r 28,000 Gap Junction Proteins from Rat Heart and Liver Are Different but Related*

(Received for publication, January 25, 1985)

Bruce J. Nicholson‡§, Daniel B. Gros¶, Stephen B. H. Kent‡, Leroy E. Hood‡, and Jean-Paul Revel‡

From the ‡Division of Biology, California Institute of Technology, Pasadena, California 91125 and the ¶Université de Poitiers, Laboratory de Biologie Cellulaire, 40 Avenue du Recteur Pineau, 86022 Poitiers Cedex, France

The sequence of the amino-terminal 32 residues of the rat heart M_r 28,000 gap junction protein presented here allows, for the first time, a sequence comparison of gap junctional proteins from different tissues (heart and liver). Comparison of the rat heart gap junction protein sequence and that available from rat liver reveals 43% sequence identity and conservative changes at an additional 25% of the positions. Both proteins exhibit a hydrophobic domain which could represent a transmembrane span of the junction. This result unequivocally demonstrates the existence of at least two forms of the gap junction protein. As yet, no homology is evident between the gap junctional proteins of either heart or liver and main intrinsic protein from rat eye lens.

Gap junctions are the mediators of direct intercellular exchanges of low molecular weight metabolites throughout the multicellular animal kingdom (1, 2). The widespread occurrence of these structures and their relatively constant and easily recognizable morphology (3) has led to the frequently expressed belief that their components might be well conserved. The consistent association of proteins of M_r 26,000–28,000 with gap junction fractions from several sources has been taken to provide support for this proposal (4–10) although some recent studies have suggested different molecular weights for the junctional protein (11–14). Interspecies comparisons of junctions from a given tissue based on immunological approaches (13, 15, 24) or by peptide map analysis (7, 16) have supported the concept of phylogenetic conservation of the protein.

In contrast, characterization of junctional proteins in different tissues has revealed a surprising diversity. It has been evident for some time that the liver gap junction protein (M_r 28,000) differs substantially from the main intrinsic protein (MIP[1]) of eye lens, an M_r 26,000 protein associated with junctional structures (17–20). However, the significance of this observation is still clouded by controversy over both the identity of the lens fiber junctions (21, 22) and the presence of MIP within these structures (15, 23). A more appropriate molecule for comparison has now become available with the isolation of the cardiac gap junction which is more characteristic in appearance than the junctions in lens. As in the case of the liver, these heart fractions contain a major protein component of M_r 28,000 (10), although recent evidence suggests that this may be derived from a larger M_r 47,000 component (12). To date, comparison of the heart junction protein with that of liver has produced apparently conflicting results. In one case, the two proteins have been claimed to be immunologically related or "even identical" (24), yet in other instances they show no cross-reactivity (25).[2] A lack of immunological relatedness would seem most compatible with the original observation that peptide maps of the two proteins show no detectable homologies (10). With no other systems available for comparison, a more reliable measure of protein homology (*i.e.* sequence analysis) for the heart and liver proteins is required to resolve the debate over the existence or degree of tissue specificity in the gap junction. The sequence analysis of the amino-terminal portion of the heart protein presented in this communication provides the first quantitative estimate of this relatedness through comparisons with the existing sequences of liver and lens (18).

MATERIALS AND METHODS

Isolation of Heart Gap Junctions—As summarized in Fig. 1, gap junctions were isolated from rat heart by a protocol essentially identical to that in Ref. 10. The final sucrose gradient used for further fractionation of the Sarkosyl-resistant material in the previous work was omitted in the present case to facilitate sequence analysis. Sixty grams (wet weight) of cardiac tissue from 75 rat hearts yielded 40 μg of M_r 28,000 protein, half of which was used for sequence analysis. The amount of protein was determined from Coomassie staining after separation by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). As a standard, we used the average value of Coomassie staining for known amounts of the Bio-Rad low molecular weight standards loaded on the same gel.

Analysis of the Final Fractions—Thin sections and negatively stained aliquots of the final fraction were prepared for examination in the electron microscope as previously described (10). After separation by SDS-PAGE (26) the protein components of these fractions were excised and analyzed by two-dimensional peptide mapping (7, 27).

Preparation of Protein and Sequence Analysis—Protein components of the Sarkosyl-resistant fraction were separated by SDS-PAGE and visualized by brief Coomassie staining and destaining. The M_r 28,000 band was then excised and electro-eluted (7), yielding approximately 200 pmol of protein. Sequence analysis by automated Edman degradation was performed in a single run on a gas-liquid phase system (28). The sample, in a volume of approximately 30 μl, was dried onto a Polybrene-coated glass filter disc under vacuum. A 30-μl aliquot of trifluoroacetic acid:H_2O (1:1, v/v) was applied to the disc and dried down as the final step. Conversion of the 2-anilino-5-

* The research was supported by Research Grant GM06965 and Biomedical Research Support Program RR07003 from the National Institutes of Health, National Cancer Institute Core Grant CA 32911, in addition to grants from Fondation Langlois, Centre National de la Recherche Scientifique (AIP Biotechnologies), Institut National de la Santé et de la Recherche Médicale (Grant 84.5.014), ANVAR (Grant X84.05.015.T), the Fondation pour la Recherche Medicale (to D. B. G.), and by an American Heart Association (Los Angeles Affiliate) Fellowship (to B. J. N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Current address: Department of Chemistry, California Institute of Technology, Pasadena, CA 91125.

[1] The abbreviations used are: MIP, main intrinsic protein; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis.

[2] O. Traub, D. Gros, and K. Willecke, unpublished observations.

thioazolinones to phenylthiohydantoins was carried out with HCl in methanol at 52 °C. The phenylthiohydantoin samples were analyzed by high pressure liquid chromatography on a cyanopropyl column (29) using a modified 21-min program. The column buffer was composed of 5% tetrahydrofuran in 20 mM Na acetate, pH 5.1. A gradient of acetonitrile without methanol was used to elute the derivatized amino acids.

RESULTS

Influence of Isolation Protocol on Sequence Analysis—Initial attempts to sequence the heart gap junction isolated by our previously published procedure (10) were unsuccessful since the amino terminus proved to be almost totally resistant to Edman degradation. Although it was possible that this blockage could have occurred *in vivo*, the successful sequencing of the liver gap junction protein (7) led us to hope that the blockage might result from modifications in the liver isolation protocol introduced in adapting it to heart. The most likely step where this could occur is the final sucrose gradient in which the junctions are exposed to 1 M urea at pH 10. These conditions could readily generate isocyanate ions which could react with primary amines such as the amino terminus of a protein, rendering it resistant to Edman degradation.

Elimination of the final sucrose gradient, the simplest rem-

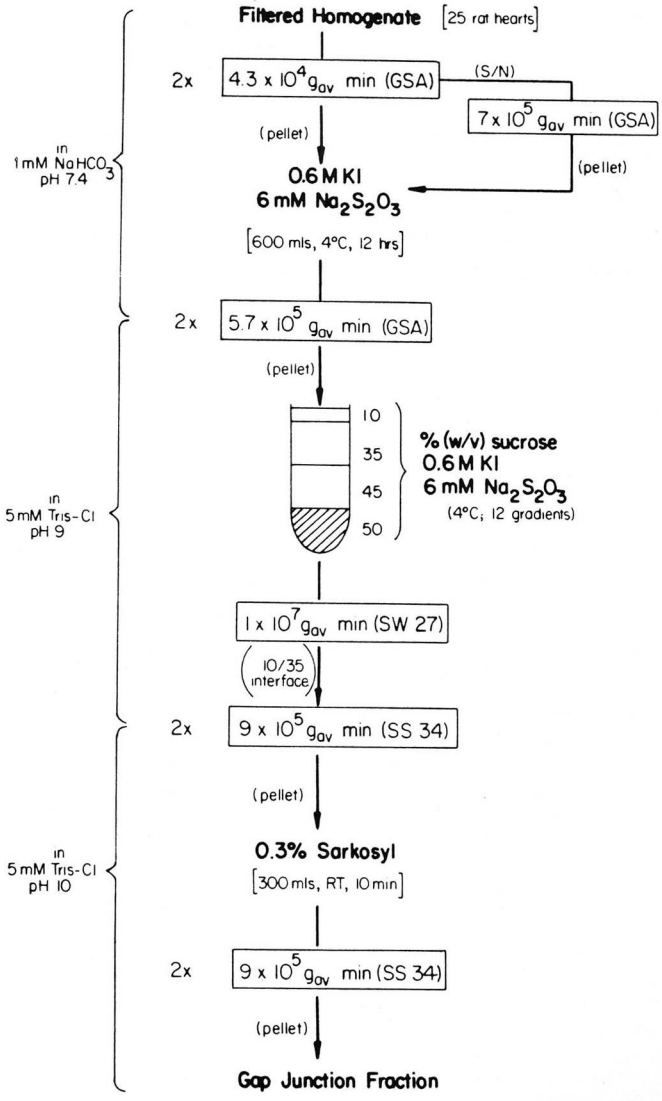

FIG. 1. **Isolation protocol for gap junctions from rat heart used in preparing material for sequence analysis.** The specific rotor used for each centrifugation is given in parentheses.

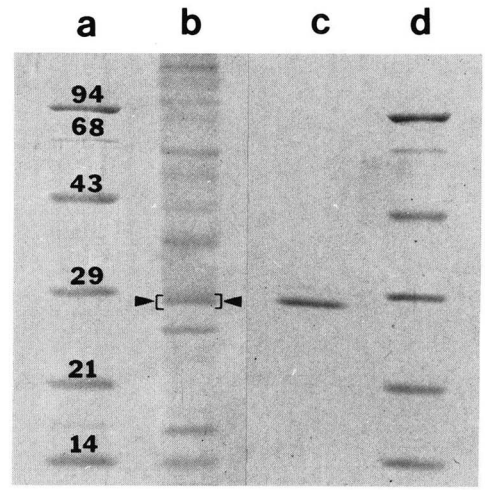

FIG. 2. **SDS-polyacrylamide gel of heart gap junction fractions.** *Lanes a* and *d*, Bio-Rad low molecular weight standards. Molecular weights are marked in thousands. *Lane b*, Sarkosyl-resistant material prepared as described in Fig. 1. The band excised for sequence analysis is bracketed. *Lane c*, heart gap junction fraction further purified by sucrose gradient as described in Gros *et al.* (10). *Lanes a* and *b* are from one gel, *lanes c* and *d* from another.

edy for this problem, causes an increase in the levels of nonjunctional material, principally single membranes and fibrous material, in the final fraction. However, gap junctions remain the most abundant structure (Fig. 1*A* in Ref. 10). Examination of these fractions by SDS-PAGE (Fig. 2, *lane b*) reveals several bands not present in fully purified junctions (Fig. 2, *lane c*). The M_r 28,000, nevertheless, is still a major component. Two-dimensional peptide mapping has been used to confirm that the M_r 28,000 protein in both the Sarkosyl-resistant fraction and fractions further purified by sucrose gradients (fully characterized in Ref. 10) are identical and that, after excision from gels, the former contains no detectable amounts of contaminating proteins (data not shown).

When cardiac gap junctions were thus isolated without exposure to the alkaline urea in the final sucrose gradient, the M_r 28,000 protein could be analyzed by automated Edman degradation. As in the case of the liver gap junctional protein (18), approximately 30 mol % of the heart junctional protein loaded on the sequenator (~200 pmol) could be detected as a unique signal in the high pressure liquid chromatography traces of the initial cleavage steps. This yield is within the range observed with other proteins analyzed in the microsequencing facility at Caltech[3] and is likely to reflect blockage of the amino terminus resulting from manipulations during isolation, frequently related to variations in the blocking activity of different batches of acrylamide. Repetitive yields for the Edman cleavage under the conditions used are 91 ± 2%, as estimated from three different residues (Leu, Ala, and Asp). As the yield of cleaved residues progressively decreases during sequencing, there is a parallel increase in signal lag (*i.e.* the residue at a given position also appears in decreasing amounts in subsequent cleavage steps). Because of this decreasing signal-to-noise ratio, the final steps in the heart sequence shown in Fig. 3 are difficult to interpret and the amino acid assignment after residue 25 should be considered tentative.

Sequence of Heart M_r 28,000 Protein—The sequence obtained for the rat cardiac gap junction protein is compared to the previously published (18) amino-terminal sequences of the rat liver gap junction protein and MIP of rat lens fiber

[3] S. Kent and M. Hunkapiller, personal communication.

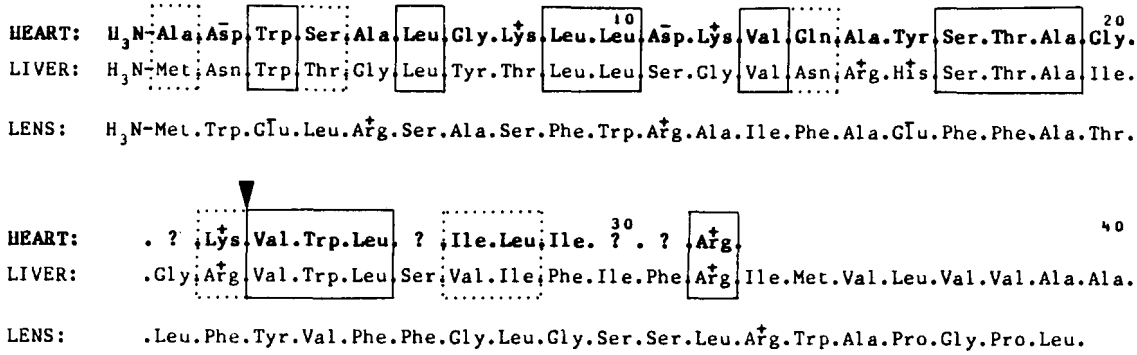

FIG. 3. **Amino-terminal sequences of the junctional proteins from rat heart, liver, and lens.** Basic (+) and acidic (−) residues are marked *above each sequence* and *question marks* denote steps in the sequences to which no amino acid could be unequivocally assigned. In the cardiac sequence, residues 26–32 are all tentative assignments due to the poor signal-to-noise ratio in this portion of the sequence. *Solid boxes* delineate residues conserved between the heart and liver proteins. *Broken boxes* indicate conservative amino acid changes in which the general properties of the residue (*i.e.* basic, acidic, polar, hydrophobic aliphatic, or hydrophobic aromatic) remain unchanged. The *arrow before residue 23* denotes the beginning of a stretch of predominantly hydrophobic residues, extending to residue 40 in the liver. Comparison of heart and liver sequences reveals 43% amino acid identity and an additional 25% of amino acid positions which show conservative substitutions. Most of the conserved residues are hydrophobic. The lens protein shares minimal homology with either heart or liver sequences, either at the amino terminus (as shown) or elsewhere in the molecule.

junctions in Fig. 3. It is readily apparent that the sequences are quite different. The heart and liver sequences, however, show 43% amino acid identity, while an additional 25% of the residues display similar properties (*i.e.* basic, acidic, polar, hydrophobic aliphatic, or hydrophobic aromatic). This alignment of residues does not require the introduction of gaps in either sequence. The overall conservation of amino acid properties at most positions is reflected in the apparent retention of a hydrophobic sequence following the basic residue at position 22 in both heart and liver proteins. Although this region has only been partially sequenced in heart, the corresponding region in liver is of just sufficient length to span the membrane. In a stretch of 18 residues, 16 of which are hydrophobic, the only charged amino acid is arginine 32 which also appears in the heart sequence (although we note again that an unambiguous assignment of residues after number 25 is difficult in the heart (see Materials and Methods)). If one allows two charged residues to be included within the lipid bilayer, the potential transmembrane region of the liver can be extended up to 23 residues. Similar situations have been described in the literature and are made thermodynamically favorable either by ionic pairing of charged residues in adjacent transmembrane spans of the protein (30) or by the location of charged or polar residues within a hydrophilic transmembrane channel (31).

In contrast, the liver and heart sequences show little or no homology with that of MIP. The 10% amino acid identity detected between lens and heart, obtained using an arbitrary alignment of the amino termini, could result from the high representation of hydrophobic amino acids in both sequences. The only features which are common to the lens protein and those of heart and liver in this alignment are 1) the amino-terminal methionine in liver and lens; 2) alanine at position 19 in all three proteins; 3) a hydrophobic, potential transmembrane region in all three sequences, although this region is located closer to the amino terminus in lens than in the other two proteins; and 4) arginine at position 32 in liver and heart and 33 in lens. Although there is no reason, *a priori*, to align the lens MIP amino terminus with those of the heart and liver proteins, alignments with other portions of the MIP sequence, deduced from the recently available cDNA clone (32), fail to reveal any more significant homologies.

DISCUSSION

The most interesting contribution of this initial sequence analysis of the heart gap junction protein is that it allows the reconciliation of data which had previously been represented as conflicting (24). The lack of detectable homology between the heart and liver proteins demonstrated by two-dimensional peptide mapping (10) and immunological studies (25),[2] is in contrast with the results of other immunological studies of tissues and Western blots which show cross-reactivity of the proteins (24). Even with the limited sequence data now available, these observations can now be seen as complementary considering the limitations of the techniques used. Differences between heart and liver proteins, particularly with regard to the distribution of charged residues, could easily account for the lack of demonstrable homology by peptide mapping either through changes in specific proteolytic cleavage sites or changes affecting the electrophoretic or chromatographic mobility of the peptides. In some cases, proteins have been found to share up to 50% amino acid identity yet produce unrelated peptide maps (33–35).

Although no potential antigenic sites are conserved between the sequences presented here, if the level of homology detected were maintained throughout the heart and liver proteins, it seems likely that some conserved regions of sufficient length to form antigenic sites (>6 amino acids (36)) should be found (statistically, there should be 1–2 such structures/250 residues). Another factor which could affect immuno-cross-reactivity is the apparently similar orientation of the heart and liver proteins in the membrane. This observation is based not only on the well conserved region of hydrophobic sequence near the amino terminus noted here, but also on the fact that proteolysis of isolated junctional plaques produces fragments of similar size in heart and liver (10). In view of all these considerations, it would seem reasonable that these proteins could show immunological cross-reactivity, but unrelated peptide maps. However, it also seems likely that only some, and not all, of the antisera produced against either the heart or liver proteins would contain antibodies against these common sites. This prediction is confirmed by comparing Refs. 24 and 25.

Since several independent lines of evidence suggest that lens MIP is isolated in the same form as synthesized on the

ribosome (32, 37), it was initially speculated that the amino-terminal methionine of liver might also be coded for by the AUG initiator codon. However, the absence of an N-terminal methionine in the heart sequence would imply that the protein isolated in this case is not identical to that originally synthesized. This raises doubts about the previous speculation, since one would have to propose that after undergoing amino-terminal, post-translational processing, the final product in the heart has a sequence which aligns exactly with an unprocessed liver protein. It would seem more likely that both are processed in a similar manner with the final cleavage occurring at a conserved site. The recent detection of higher molecular weight proteins in both the liver (M_r 54,000 (14)) and heart (M_r 47,000 (12)), which seem related to the more generally seen M_r 28,000 polypeptides and are highly labile to proteolysis, suggests that post-translational processing of both proteins may occur.

The most surprising finding of this study is the extent of divergence of these two proteins which form morphologically very similar structures in heart and liver. It is possible that this diversity reflects a lack of selective pressure to conserve specific residues, with only general features being retained. However, it is also possible that some of the variation may relate to differences in the function of gap junctions from tissue to tissue. In this light, it is of interest to note that, to date, heterogeneity of the gap junctional protein between species (15, 16, 18) has been much less than that between tissues (10, 18, 25). Regions of the protein conserved between both species and tissues would, on the other hand, reflect universal features of the gap junction. The presence of a single basic residue (arginine 32) within a strongly hydrophobic region of both liver and heart proteins could represent such a conserved site, with possible relevance to the channel structure of gap junctions.

Although lens MIP is clearly quite different from the gap junctional proteins from other sources sequenced here, a more complete analysis of the extent of gap junction diversity will be needed before the nature of MIP can be determined. This study represents an initial step in defining the specific structural variations between junctions in different tissues.

Acknowledgments—Our thanks go to Dr. Barbara Yancey and Jan Hoh for constructive discussions and to the Biology Division's secretarial staff for typing the manuscript.

REFERENCES

1. Lowenstein, W. R. (1981) *Physiol. Rev.* **61,** 829–911
2. Bennett, M. V. L. (1978) in *Intercellular Junctions and Synapses* (Feldman, J., Gilula, N. B., and Pitts, J. D., eds) pp. 25–36, Chapman and Hall, London
3. Staehelin (1974) *Int. Rev. Cytol.* **39,** 191–283
4. Duguid, J., and Revel, J.-P. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **40,** 45–47
5. Henderson, D., Eibel, H., and Weber, K. (1979) *J. Mol. Biol.* **132,** 193–218
6. Hertzberg, E. L. (1980) *In Vitro* (*Rockville*) **16,** 1057–1067
7. Nicholson, B. J., Hunkapiller, M. W., Grim, L. B., Hood, L. E., and Revel, J.-P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 7594–7598
8. Broekhuyse, R. M., Kuhlman, E. D., and Stols, A. L. H. (1976) *Exp. Eye Res.* **23,** 365–371
9. Goodenough, D. A. (1979) *Invest. Ophthalmol. Visual Sci.* **18,** 1104–1122
10. Gros, D. B., Nicholson, B. J., and Revel, J.-P. (1983) *Cell* **35,** 539–549
11. Finbow, M. E., Shuttleworth, J., Hamilton, A. E., and Pitts, J. D. (1983) *EMBO J.* **2,** 1479–1486
12. Manjunath, C. K., Goings, G. E., and Page, E. (1984) *Am. J. Physiol.* **246,** H865–H875
13. Finbow, M. E., Eldridge, T., Buultjens, J., Lane, N. J., Shuttleworth, J., and Pitts, J. D. (1984) *EMBO J.* **3,** 2271–2278
14. Warner, A. E., Guthrie, S. C., and Gilula, N. B. (1984) *Nature* **311,** 127–131
15. Bok, D., Dockstader, J., and Horwitz, J. (1982) *J. Cell Biol.* **92,** 213–220
16. Takemoto, L. J., Hansen, J. S., and Horwitz, J. (1981) *Comp. Biochem. Physiol.* **68B,** 101–106
17. Hertzberg, E. L., Anderson, D. J., Friedlander, M., and Gilula, N. B. (1982) *J. Cell Biol.* **92,** 53–59
18. Nicholson, B. J., Takemoto, L. J., Hunkapiller, M. W., Hood, L. E., and Revel, J.-P. (1983) *Cell* **32,** 967–978
19. Ziegler, S. J., and Horwitz, J. (1981) *Invest. Ophthalmol. Visual Sci.* **21,** 46–51
20. Traub, O., and Willecke, K. (1982) *Biochem. Biophys. Res. Commun.* **109,** 895–901
21. Kuszak, J. R., Rae, J. L., Pauli, B. V., and Weinstein, R. S. (1982) *J. Ultrastruct. Res.* **81,** 249–256
22. Zampighi, G., Simon, S. A., Robertson, J. D., McIntosh, T. J., and Costello, M. J. (1982) *J. Cell Biol.* **93,** 175–189
23. Paul, D. L., and Goodenough, D. A. (1983) *J. Cell Biol.* **96,** 625–632
24. Hertzberg, E. L., and Skibbens, R. V. (1984) *Cell* **39,** 61–69
25. Dermietzel, R., Leibstein, A., Frixen, U., Janssen-Timmen, U., Traub, O., and Willecke, K. (1984) *EMBO J.* **3,** 2261–2270
26. Laemmli, U. K. (1970) *Nature* **227,** 680–686
27. Elder, J. H., Pickett, R. A., Hampton, J., and Lerner, R. A. (1977) *J. Biol. Chem.* **252,** 6510–6515
28. Hewick, R. M., Hunkapiller, M. W., Hood, L. E., and Dreyer, J. W. (1981) *J. Biol. Chem.* **256,** 7990–7997
29. Hunkapiller, M. W., and Hood, L. E. (1983) *Methods Enzymol.* **91,** 486–493
30. Engelman, D. M., Henderson, R., McLachlan, A. D., and Wallace, B. A. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 2023–2027
31. Finer-Moore, J., and Stroud, R. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 155–159
32. Gorin, M. B., Yancey, S. B., Cline, J., Revel, J.-P., and Horwitz, J. (1984) *Cell* **39,** 49–59
33. Stevens, F. C., Walsh, M., Ho, H. C., Leo, T. S., and Wang, J. H. (1976) *J. Biol. Chem.* **251,** 4495–4500
34. Watterson, D. M., Sharief, F., and Vanaman, T. C. (1980) *J. Biol. Chem.* **255,** 962–975
35. Hood, L., Grant, J. A., and Sox, H. C., Jr. (1970) in *Development Aspects of Antibody Formation and Structure* (Sterzl, J., ed) Vol. 1, pp. 283–309, Academia Publishing House of Czechoslovak Academy of Sciences, Prague
36. Atassi, M. Z. (1975) *Immunochemistry* **12,** 423–438
37. Paul, D. L., and Goodenough, D. A. (1983) *J. Cell Biol.* **96,** 636–638