Extensive Homology Between the Subunits of the Phytohemagglutinin Mitogenic Proteins Derived from Phaseolus vulgaris

(glycoprotein/lymphocyte)

J. BRUCE MILLER, ROBERT HSU, ROBERT HEINRIKSON, AND STANLEY YACHNIN

Departments of Medicine and Biochemistry, The University of Chicago, and the Franklin McLean Memorial Research Institute (operated by the University of Chicago for the United States Atomic Energy Commission), Chicago, Illinois 60637

Communicated by Leon O. Jacobson, January 23, 1975

ABSTRACT  The phytohemagglutinin mitogenic proteins derived from Phaseolus vulgaris comprise a class of five glycoproteins that are isomeric tetramers composed of varying proportions of two different subunits (L and R). Within the native tetramer, the L subunit is a potent leukoagglutinin and mitogen that lacks hemagglutinating properties, whereas the R subunit is a potent hemagglutinin with little or no mitogenic activity. The subunits have been isolated in homogeneous form by isoelectric focusing in 8 M urea. Previous work has shown that they have equal molecular weights and differ in amino-acid sequence from residues 1-7, but are identical in positions 8-24 [1973 J. Exp. Med. 138, 939-951]. We now report amino-acid composition studies which reveal striking similarities between the subunits. Both lack methionine and cysteine. The twelfth residue in each subunit is a glycosylated asparagine, with the identical carbohydrate composition in each. The last three residues of the subunits, as determined by carboxypeptidase A digestion, are identical. Tryptic peptide mapping of the succinylated phytohemagglutinin subunits reveals a high degree of similarity. We conclude that the substantial difference in biological properties among the tetrameric phytohemagglutinin mitogens is a result of relatively restricted differences in the primary structure of their constituent subunits.

The phytohemagglutinin (PHAP) mitogenic proteins derived from the red kidney bean, Phaseolus vulgaris, have been shown to comprise a family of five heterogeneous proteins (1, 2). They consist of isomeric, noncovalently bound tetramers which are made up of two different subunits, designated as L and R (Fig. 1) (2-4). One of the five proteins is a potent leukoagglutinin with low hemagglutinating activity (L-PHAP); it is homogeneous, consisting of four identical subunits (1). Three other, closely related proteins, which have modest leukoagglutinating but potent hemagglutinating properties (H-PHAP), have also been isolated. They consist of hybrid tetramers containing varying proportions of the two subunits (2L-2R, 1L-3R, and 4R), the increasing R subunit content of which is reflected by increasing cathodal migration on polyacrylamide gel electrophoresis. A fifth PHAP mitogenic protein (3L-1R) has been identified, but detailed study of its properties has been hampered in the past by its contamination with other proteins. We have tentatively concluded that, within each native tetramer, the L subunit has strong mitogenic activity and a high affinity for receptors of lymphocyte membranes, but little or no affinity for those of erythrocytes.

Conversely, the R subunit has a high affinity for erythrocyte membrane receptors, but little for those of lymphocytes. As a result, the 4R tetramer displays little or no lymphocyte mitogenic activity (6). The hybrid molecules (1L-3R, 2L-2R, and 3L-1R tetramers) have been found to be mitogenic, to cause mixed agglutination of erythrocytes and lymphocytes, and to have enhanced lymphocyte-transforming capability in the presence of erythrocytes (1-5).

The molecular weights of the L and R subunits, as determined by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (8), have been shown to be identical (4). It is possible to separate the subunits by isoelectric focusing in 8 M urea. Partial sequence analysis of the two subunits has shown that, although they differ in six of the first seven residues from the NH2-terminus, they are identical in sequence from residue 8 through residue 24 (4). This paper presents further evidence supporting our hypothesis of extensive structural homology between the two PHAP subunits.

MATERIALS AND METHODS

The isolation and purification of tetrameric L- and H-PHAP from commercial Bacto-Phytohemagglutinin P (Difco Laboratories), and the isolation of pure L and R subunits by isoelectric focusing in 8 M urea of L- and H-PHAP, have been described (1, 4).

Amino-acid analyses of the L and R subunits were performed by hydrolysis of the salt-free protein in 4 M methanesulfonic acid that contains 0.2% 3-(2-amino-ethyl)indole at 115° for 24, 48, and 72 hr, followed by addition of equal amounts of 3.5 M NaOH, as described by Liu (9). Analysis was performed on a Durrum D-500 analyzer. This method allows direct quantitation of tryptophan content by amino-acid analysis.

Determination of the NH2-terminal sequences of the purified L and R subunits by automated Edman degradation with a Beckman protein-peptide sequenator (model 890-B; Beckman Instruments, Inc.) has been described in an earlier paper (4). The amino-acid and sugar components of the twelfth residue in the L and R subunits (200-230 nmol) were determined by a modification of the automated Edman degradation method at cycle 12, in order to permit extraction and isolation of the phenylthiohydantoin (PhNCS). This was presumed to contain a sugar moiety in covalent form (4). After coupling in Quadrol, cyclization, and cleavage in anhydrous heptafluorobutyrate, the residual sugar-PhNCS-amino acid was extracted manually in 50% aqueous ethanol and evapo.
rated to dryness. After addition of mannitol (0.04–0.2 μmol) to an aliquot of 80 nmol, the residual sugar-PhNCS was converted to methylglycosides by methanolation in 1.0 M HCl for 16 hr at 80° (10, 11), neutralized with silver carbonate, and again N-acetylated with acetic anhydride for 16–24 hr (12). Component sugars were quantitated and identified in a single procedure by gas-liquid chromatography of the trimethylsilylated sugars, as described by Dawson et al. (13). For identification and quantitation of the amino-acid moiety of residue 12, 50 nmol of PhNCS-norleucine were added as an internal standard to an aliquot of 40 nmol; the sample was then dried for 15 min and hydrolysed under reduced pressure with 57% HI at 130° for 24 hr (14). In a separate experiment, 50 nmol each of aspartate and N-acetylglucosamine were quantitated on an amino-acid analyzer after hydrolysis in sealed, evacuated tubes, either with 4 M HCl for 3 hr at 100° (15) or with 4 M methanesulfonic acid containing 0.2% trypthamine for 12 hr at 100° (16). The efficiency of the cleavage of phenylhydantoin from PhNCS-amino acid and the destruction of sugars during hydrolysis were measured by identical treatment of PhNCS-aspartate and amino sugar standards. A control experiment with 1-L-aspartamido(2-acetamido)-1,2-dideoxy-β-D-glucose (Asn-GlcNAc) was run in parallel with the unknown.

**Table 1. Amino-acid composition of L and R subunits**

<table>
<thead>
<tr>
<th>Residues per 100 amino-acid residues</th>
<th>Observed residues per subunit*</th>
<th>Residues per subunit extrapolated to T1</th>
<th>Mole integer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L</td>
</tr>
<tr>
<td>Aspartic acid and asparagine</td>
<td>15.24</td>
<td>14.54</td>
<td>47.98</td>
</tr>
<tr>
<td>Threonine</td>
<td>8.06</td>
<td>9.19</td>
<td>25.37</td>
</tr>
<tr>
<td>Serine</td>
<td>10.65</td>
<td>11.10</td>
<td>33.53</td>
</tr>
<tr>
<td>Glutamic acid and glutamine</td>
<td>7.76</td>
<td>6.34</td>
<td>24.43</td>
</tr>
<tr>
<td>Proline</td>
<td>1.71</td>
<td>1.83</td>
<td>5.38</td>
</tr>
<tr>
<td>Glycine</td>
<td>8.50</td>
<td>7.25</td>
<td>26.78</td>
</tr>
<tr>
<td>Alanine</td>
<td>7.27</td>
<td>7.38</td>
<td>22.88</td>
</tr>
<tr>
<td>Half-cystine</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Valine</td>
<td>8.59</td>
<td>8.10</td>
<td>27.04</td>
</tr>
<tr>
<td>Methionine</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>4.83</td>
<td>5.00</td>
<td>15.21</td>
</tr>
<tr>
<td>Leucine</td>
<td>8.87</td>
<td>9.71</td>
<td>27.57</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>2.16</td>
<td>2.38</td>
<td>6.80</td>
</tr>
<tr>
<td>Phenylalanine‡</td>
<td>7.07</td>
<td>6.79</td>
<td>22.25</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.85</td>
<td>1.15</td>
<td>2.67</td>
</tr>
<tr>
<td>Lysine</td>
<td>4.29</td>
<td>5.03</td>
<td>13.52</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>0.92</td>
<td>1.13</td>
<td>2.90</td>
</tr>
<tr>
<td>Arginine</td>
<td>3.38</td>
<td>2.10</td>
<td>10.59</td>
</tr>
</tbody>
</table>

* Calculated on the basis of subunit molecular weight of 34,000 (4).  
† Hydrolyses were carried out for 24, 48, and 72 hr, and the results were back-extrapolated to time T4 to correct for decomposition.  
‡ Sum of phenylalanine plus glucosamine and galactosamine (see text).  

**Fig. 1.** Schematic representation of the tetrameric structure of the five phytohemagglutinin mitogens, consisting of varying proportions of the L and R subunits.

**Fig. 2.** Results of sequential amino-acid analysis of the L and R subunits. Note: The residues in italics (residues 8–24) are shown here only once since they are identical for both L and R subunits.

Comparison of the carboxyl-terminal sequences in the L and R subunits was based upon analysis of the time course of release of amino acids during digestion with carboxypeptidase A (Worthington Biochemical Corp.). After succinylation of amino groups, a solution of the protein (8 mg/ml) in 0.2 M N-ethylmorpholine acetate (pH 8.0) was incubated at 1° with 1% by weight of carboxypeptidase A, and aliquots were removed at 1, 4, 8, and 12 min.

Digestion of succinylated L- and H-PHAP with 1-L-tosyl-amido-2-phenylethyl chloromethyl ketone-treated trypsin (Worthington Biochemical Corp.) was carried out at an enzyme-to-substrate ratio of 1:100 in 0.2 M N-ethylmorpholine acetate buffer (pH 8.0) at 37° for 4 hr, with a second addition of trypsin at 2 hr. Tryptic peptide maps were obtained by high-voltage paper electrophoresis at pH 6.5 (pyridine/acetic acid/...
RESULTS

Amino-Acid Analysis of L and R Subunits. Pure L and 95% pure R subunits obtained from isoelectric focusing of L- and H-PHAP, respectively, were hydrolyzed and analyzed. The results are given in Table 1. The compositions were based on the subunit molecular weight of 34,000 (4, 18, 19). Considerable similarity in the composition of the two subunits is evident. No methionine or cysteine was found in either protein, and both contained a predominance of aspartate, serine, threonine, and leucine. Although the total number of arginine plus lysine residues (23) was the same in each subunit, there were 10 arginine residues in the L as opposed to seven in the R subunit. Three tryptophan residues per subunit were found for both L and R subunits. We verified that the amino sugars glucosamine and galactosamine appear quantitatively superimposed on the phenylalanine peak obtained from the Durrum D-500 analyzer. Therefore, the values obtained for phenylalanine consist of the total of phenylalanine plus glucosamine and galactosamine.

Identification of Asparagine-12 as a Site of Glycosylation. The results of amino-terminal sequence analysis of the L and R subunits, as reported previously, are shown in Fig. 2. After repeated failure, with both subunits, to recover an amino acid in the twelfth position, we speculated that this might be a site of glycosylation (4). Therefore, we halted the automated Edman degradation procedure at the twelfth cycle before extraction, and the reaction chamber contents were manually extracted and analyzed. These studies revealed the presence of aspartic acid and glucosamine in a ratio of 1:2. Accordingly, the original amino acid at position 12 is assumed to be asparagine (20). Carbohydrate analysis showed a sugar moiety consisting of two glucosamine and six to seven mannose residues in each subunit.

COOH-Terminal Sequence. Results of carboxypeptidase A digestion, at 1°, of succinylated L-PHAP (pure L subunit) and H-PHAP (R-to-L subunit ratio of 4:1) are shown in Fig. 3. For both subunits there was an initial release of leucine, followed by isoleucine and then serine. No other residues were found. We conclude that the L and R subunits terminate in the same sequence, i.e., -Ser-Ile-Leu-COOH.

Tryptic Peptide Mapping. Composite tryptic peptide maps obtained from the results of over 15 determinations on digests from succinylated L- and H-PHAP are presented in Fig. 4. A total of eleven peptides was found for L-PHAP and nine for H-PHAP. Five peptides, indicated by the shaded spots in Fig. 4, appear to be identical. Since the COOH-terminal tryptic peptide from each subunit should be the only fragment not containing arginine, it should be the only ninhydrin-positive peptide that gives no fluorescence in the arginine test. The single blackened spots on each map meet these criteria and, therefore, presumably represent the COOH-terminal peptides from the L and R subunits. Their identical rates of migration during electrophoresis and chromatography suggest that they are identical.
DISCUSSION

We have previously reported the homology between the PHAP subunit structures. The estimated molecular weight of the L and R subunits, as determined by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (8), was 34,000 (4). They were identical in sequence at the NH₂-terminus from residue 8 through 24 (4), even though they differed in six of the initial seven residues. Because we were unable to identify the twelfth residue from an automated Edman degradation of both the L and R subunits, we hypothesized that this might be a glycosylated residue. In several glycoproteins, glycosylated asparagine residues have been shown to occur in a triplet sequence in which they are separated by one amino acid from a serine or threonine on their COOH-terminal side (20). The twelfth residue of each subunit satisfies these criteria. In such a glycopeptide, the aspartic acid is linked directly to N-acetylglucosamine in the form of 1-L-β-aspartamido(2-acetamido)-1,2-dideoxy-β-D-glucose. Such a carbohydrate-protein linkage was first identified in ovalbumin (21, 22), and it appears now to occur in many animal proteins (23). Its presence in plant glycoproteins was first described in soybean hemagglutinin (24) and more recently in bromelain and other plant glycoproteins (25). Our present results indicate that residue 12 of both subunits is a glycosylated asparagine. Moreover, the compositions of the carbohydrate moieties also proved to be nearly identical, consisting of two glucosamine and six to seven mannose residues. These were the only sugar components of crystalline kidney-bean leucoagglutinin reported by Räsänen et al. (16).

The amino-acid composition data for the L and R subunits reported here reveal the complete absence of any methionine or half-cystine residues. This is consistent with the analysis, by Räsänen et al. (16), of kidney-bean leucoagglutinin, in which no sulfur-containing amino acids were noted. The total number of arginine plus lysine residues is identical for the L and R subunits; however, there are ten and seven residues of arginine, respectively, in the two polypeptides. One would, therefore, predict that trypsin digestion of the succinylated subunits in which lysine residues are blocked would result in 11 peptides for L-PHAP and eight for H-PHAP. The observed values of 11 and nine are consistent with this prediction.

The tryptic peptide maps of succinylated L- and H-PHAP reported here show a high degree of similarity, with six peptides appearing to be identical. Of special interest is the apparent identity between the carboxyl-terminal tryptic peptides of L- and H-PHAP. This is in agreement with our finding of identical sequences for the last three amino acid residues of L- and H-PHAP.

We conclude that the very different biological properties of the L and R subunits of the phytohemagglutinin mitogenic proteins must arise from relatively restricted differences in their primary structures. These biological differences may stem from amino acid sequence differences in close proximity to the carbohydrate at position 12, or in some as yet undetermined position of the subunits.

We gratefully acknowledge the assistance of Drs. Claudia Noyes and Glyn Dawson. This work was supported in part by grants from NSF (GB 29098), USPHS (5-T01 AM-05620), and NCI (CA-05230).