

Identification of the Cystines Which Link the Acidic and Basic Components of the Glycinin Subunits*

(Received for publication, December 19, 1983)

Paul E. Staswick, Mark A. Hermodson, and Niels C. Nielsen

From the United States Department of Agriculture and the Departments of Agronomy and Biochemistry, Purdue University, West Lafayette, Indiana 47907

The half-cystine residues involved in linking the acidic and basic polypeptides were determined for several glycinin subunits. The cystines were localized with specific cyanogen bromide fragments either by comparing the electrophoretic mobility of nonreduced and reduced fragments, or by co-purifying and then determining the NH₂-terminal sequence of the covalently linked fragments. Residues involved in the disulfides were further identified by labeling them with [³H]iodoacetic acid. Only 1 cystine was found to be involved in linking the acidic and basic components of each subunit, and they were in analogous positions in each of the subunits studied. Potential sites for intrapolypeptide cystines were also identified.

Glycinin from soybean (*Glycine max* L. Merr.) cultivar CX635-1-1-1 contains at least five major kinds of subunits, each subunit being composed of an acidic and a basic polypeptide. With the exception of one polypeptide (A₄), the acidic and basic components are covalently linked by disulfide bonds (1). The covalently linked subunits are denoted A_{1a}B₂, A_{1b}B_{1b}, A₂B_{1a}, A₃B₄, and A₅B₃. Previous studies have revealed considerable homology among the acidic and among the basic polypeptides, although distinct differences were evident (2, 3).

Since the complete amino acid sequence for the A₂B_{1a} subunit is known (4, 5), a study was undertaken to identify the cystine(s) linking the two subunit components. Evidence for only one intermolecular cystine was found in this subunit. A single disulfide bond at an analogous position was also found for each of the other covalently linked subunits.

MATERIALS AND METHODS

Protein Purification—Unreduced glycinin was extracted from dry soybean seeds of cultivar CX635-1-1-1, and all free sulfhydryl groups were S-alkylated with 4-vinylpyridine in 6 M guanidine hydrochloride as described previously (1). The subunits were then purified by DEAE-Sephadex chromatography, dialyzed against 9% formic acid, and lyophilized.

CNBr Digestion and Fragment Analysis—Each nonreduced subunit (10–20 mg) was cleaved with CNBr (4), and the fragments were analyzed on a 25% polyacrylamide SDS¹ gel by a modification of the

Laemmli system (6, 7). A portion of each digest was reduced with 2-mercaptoethanol (2%) and electrophoresed next to the unreduced sample. The rest of each cleaved sample was fractionated on a Sephadex G-75 column (2.5 × 100 cm) in 9% formic acid. Some fractions from the G-75 column were further purified by reverse-phase HPLC (4). The NH₂-terminal sequences of the purified peptides were determined on a Beckman 890C Sequencer. Cyanogen bromide fragments were named according to their expected order in the respective molecules (*i.e.* A_{1a}-CB₂, B_{1a}-CB₁) (4, 5).

Radiolabeling and Analysis of Cystine Residues—Purified disulfide-linked peptides were dissolved in 0.5 ml of 6 M guanidine HCl buffer containing 0.13 M Tris-HCl (pH 7.6) and EDTA (1 mg/ml). Dithiothreitol was added to a 10-fold molar excess over the estimated cystine content, and the mixture was kept in the dark at room temperature. After 2 h the reduced peptides were desalted on a Sephadex G-25 column (1 × 17 cm) in 9% formic acid and lyophilized. The lyophilized material was redissolved in the same buffer and [³H]iodoacetic acid (100,000 cpm) was added. After 10 min in the dark, a quantity of iodoacetamide sufficient to react with all remaining free sulfhydryl groups was introduced into the reaction. This mixture was desalted after 10 min as described above. The radiolabeled peptide was sequenced, and the material from each cycle was prepared for analysis in the normal way (4). One-fourth (50 μl) of the sample from each cycle was analyzed for tritium in a Beckman LS 7500 Liquid Scintillation Counter.

RESULTS²

The initial strategy used to determine the position of disulfide bonds linking the acidic and basic polypeptides in the subunits of glycinin involved a comparison of the relative mobilities on SDS gels of reduced and nonreduced cyanogen bromide-generated peptides from each subunit. An increase in mobility following reduction indicated that the fragment was involved in a disulfide bond. The number of fragments expected from each polypeptide could be predicted, since the methionine content of each acidic and basic polypeptide had been determined previously (2). (Since the total number of residues was overestimated for A₂ (4) and presumably for the other acidic polypeptides as well, the values given for them should be reduced by about 18%.) The NH₂-terminal sequences and the relative positions of most CNBr fragments have also been determined for the acidic polypeptides (3).

The electrophoretic analysis of subunit A₃B₄ is shown in Fig. 1. Since B₄ contains only 1 methionine and it is the ninth residue from the NH₂ terminus (2), cleavage of the polypeptide by CNBr should yield two fragments that have *M_r* ≈ 20,000 and *M_r* ≈ 1,000. The darkest band in both the reduced and nonreduced lanes for A₃B₄ corresponded to the large fragment. Since its relative mobility did not change in the presence of mercaptoethanol, it was not involved in an interchain disulfide bond. A₃ yielded three fragments (*M_r* ≈ 22,000, *M_r* ≈

* This work was supported by the United States Department of Agriculture Competitive Grants Program and the American Soybean Association Research Foundation. This is cooperative research of the United States Department of Agriculture-Agricultural Research Service and the Purdue Agricultural Experimental Station and is Journal Paper 9716. 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.

¹ The abbreviations used are: SDS, sodium dodecyl sulfate; HPLC, high-pressure liquid chromatography.

² Primary data from the sequenator for cyanogen bromide fragments described in this paper are retained in our files and will be made available upon request.

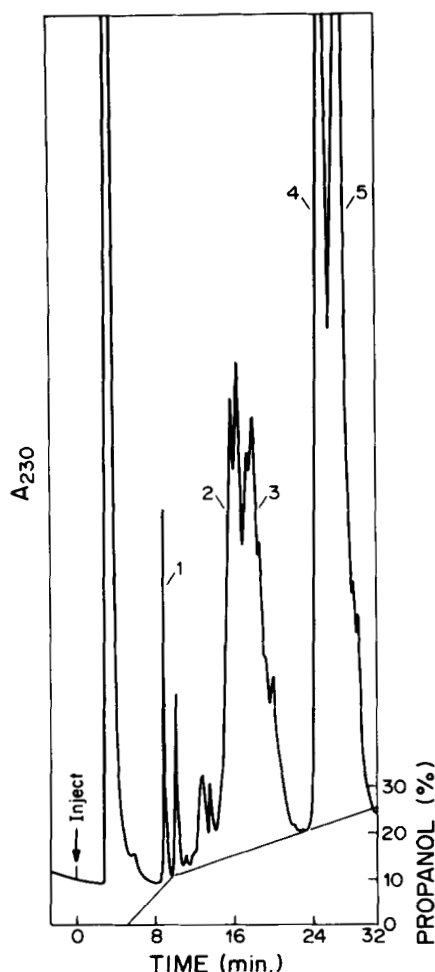


FIG. 4. HPLC separation of the A_2B_{1a} -F₄ peptides. Lyophilized protein from F₄, Fig. 3, was subjected to reverse-phase HPLC as described above. Protein from each peak was identified by NH_2 -terminal sequence analysis. Peak 3 contained the disulfide-linked fragments A_2 -CB-2 and B_{1a} -CB-1. Other peaks were: 1) A_2 -CB-6; 2) B_{1a} -CB-3; 4) A_2 -CB-5; 5) B_{1a} -CB-2.

identified by their NH_2 -terminal sequences, each cycle of degradation of peak three was found to contain two phenylthiohydantoin derivatives in equimolar amounts, which indicated that two polypeptides were present. The sequences corresponded to the second cyanogen bromide fragment from A_2 (A_2 -CB-2) (4) and the first fragment from B_{1a} (B_{1a} -CB-1) (5). Since B_{1a} -CB-1 contained only 9 residues, it was not expected to co-purify with A_2 -CB-2 (51 residues) unless the two were covalently linked. Both fragments contained cysteine residues which were shown to be involved in the disulfide bond joining them (see below). The only other fragments from this subunit which contained cysteine residues potentially involved in intermolecular disulfide bonds were A_2 -CB-1, A_2 -CB-6, and B_{1a} -CB-3. These peptides were each found unassociated with any other peptide (Figs. 3 and 4). Thus, it was apparent that only one intermolecular disulfide bond covalently linked the two components of the A_2B_{1a} subunit of glycinin.

Although the complete amino acid sequence was not available for the $A_{1a}B_2$ subunit, its high degree of homology with A_2B_{1a} permitted it to be analyzed in a similar manner. The Sephadex G-75 fraction analogous to F₄ (Fig. 3) was purified by HPLC (data not shown) and was found to contain two peptides, A_{1a} -CB-2 and B_2 -CB-1, which co-purified in equimolar amounts. Again because B_2 -CB-1 has only 10 residues

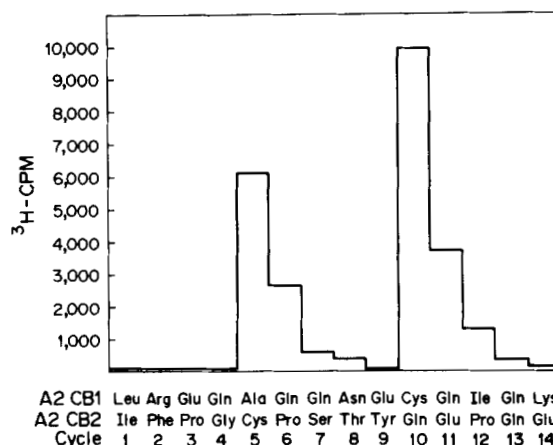


FIG. 5. [^3H]iodoacetic acid modification of the A_2 -CB-1 fragment. A_2 -CB-1 was mixed with A_2 -CB-2 at a ratio of 3:2 and labeled as described. Label was detected in cycle 10 (A_2 -CB-1) as well as in cycle 5 (A_2 -CB-2).

and A_{1a} -CB-2 has over 40, these fragments would not be expected to co-chromatograph unless they were joined by the intermolecular disulfide bond linking A_{1a} to B_2 .

Whereas the small cyanogen bromide fragments from B_{1a} , B_2 , and B_4 each contained only 1 cysteine, the respective acidic fragments each contained at least two potential sites which could be involved in the disulfide bond between fragments. The specific residues involved in this covalent linkage was determined for A_{1a} -CB-2, A_2 -CB-2, and A_3 -CB-3 by radiolabeling the cysteine residues with [^3H]iodoacetic acid after cleavage of the disulfides with dithiothreitol. All free sulfhydryl groups had previously been modified with 4-vinylpyridine and thus could not react with the iodoacetic acid. When the A_{1a} -CB-2 and A_2 -CB-2 fragments were sequenced and the products analyzed for radioactivity, the cysteine in cycle five of each was found to be heavily labeled, but no label was found in the other Cys residue at position 18 in each fragment. In both analyses some label was found at position 7 due to labeling of the Cys in the 9-residue fragment from B_2 and B_{1a} , respectively. The small size of these fragments resulted in losses from the sequenator cup and losses during desalting procedures and only one-third as many counts were recovered at these positions as were recovered at position 5 in each degradation. This was consistent with the observed yields of the phenylthiohydantoin derivatives from the respective fragments.

The NH_2 -terminal fragment (CB-1) of A_2 contained 2 cysteine residues at positions 10 and 51 (4) which could be joined by a disulfide bond. This possibility was further studied by labeling purified A_2 -CB-1 with [^3H]iodoacetate after reduction with dithiothreitol (as above). A_2 -CB-2 was also added to the sample as an internal standard, and the ratio of CB-1 to CB-2 was about 3:2, as determined by the yields of each from the sequencer. The cysteine in CB-1 (cycle 10) was labeled to the same relative extent as the one in CB-2 (Fig. 5) and suggested that it could be involved in a disulfide bond with the cysteine at position 51, the only other cysteine in this fragment.

The cysteine-containing fragment from A_3B_4 was also purified and analyzed by procedures similar to those employed for the other subunits. Only cysteine 17 of the second cyanogen bromide fragment of the acidic chain was labeled to a significant extent. This is in the homologous position to the Cys residues of the other acidic polypeptides found to be joined to the basic polypeptides (3). A summary of the data is presented in Fig. 6.

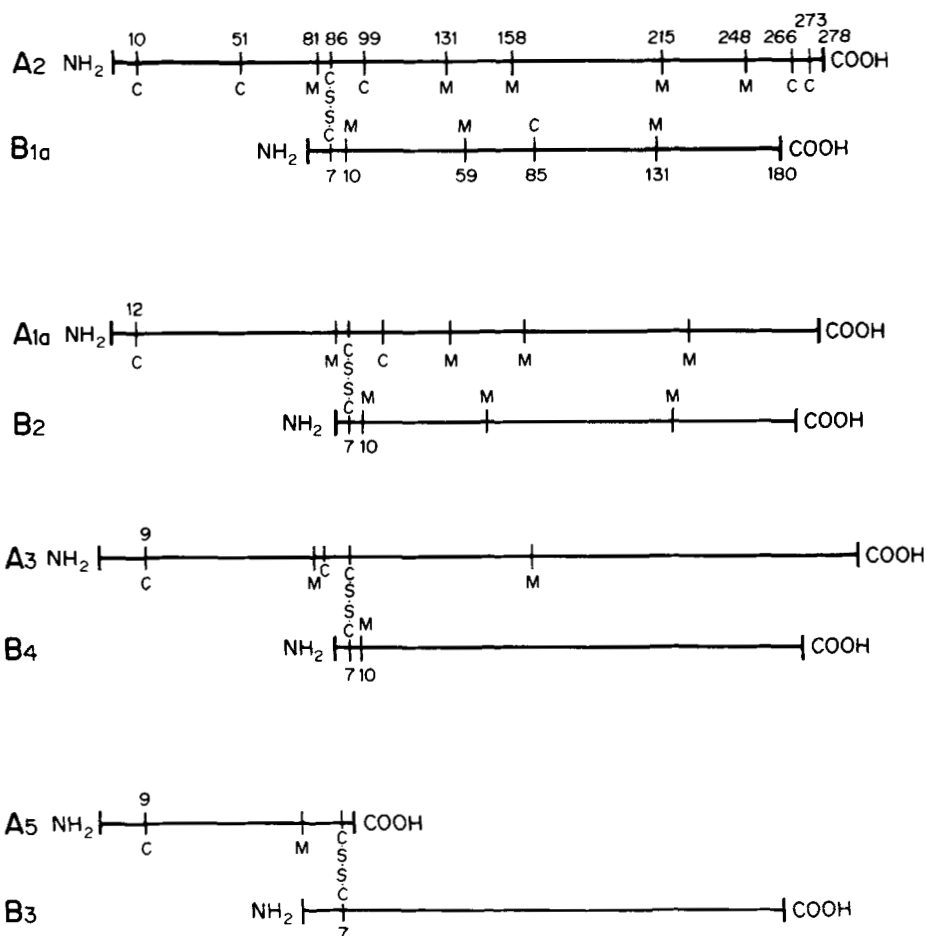


FIG. 6. Relative position of the interchain disulfide bonds in the glycinin subunits. Horizontal lines represent polypeptides and are proportional in length to the estimated length of the protein. The residue number of methionines (M) and cysteines (C) are indicated from the complete sequence for A_2B_{1a} . The position of these residues in other subunits is estimated from the size of CNBr fragments and sequence homology with A_2B_{1a} . Additional cysteine residues are undoubtedly present in $A_{1a}B_2$, A_3B_4 , and A_5B_3 but their positions are not known. Interchain disulfides are indicated by -C-S-S-C-.

DISCUSSION

In an earlier report we identified which acidic and basic polypeptides of the glycinin subunits were linked via disulfide bonds (1). The present study was undertaken to determine the precise location of those interchain bonds. Our results demonstrate that only one intermolecular disulfide bond is involved in linking the acidic and basic polypeptides. Similar bonds are found at homologous positions in each subunit. In the case of A_2B_{1a} where the exact residue positions are known, Cys 7 in the basic component is joined to Cys 86 in A_2 (Fig. 6). The location of the intermolecular disulfide bond was not determined for the $A_{1a}B_{1b}$ subunit. However, the high degree of homology that the polypeptides in this subunit have with the acidic and basic ones in the other subunits make it likely that the bond will occur in exactly the same place there as well. This information further emphasizes the essentially similar nature of the subunits of which glycinin is composed.

The possibility that small cyanogen bromide fragments containing another intermolecular disulfide bond had gone undetected was ruled out by examining the complete amino acid sequence for A_2B_{1a} . Nor is it likely that such fragments were present in other subunits either, since their amino acid compositions indicated they contained fewer methionine residues than A_2B_{1a} , and fewer fragments should have resulted when these subunits were cleaved with cyanogen bromide.

Assuming that all cysteine residues not involved in disulfide bonds had previously been modified with 4-vinylpyridine, the incorporation of iodoacetic acid into cysteine 10 of A_2 implicated it as a half-cystine residue. The other half-cystine in A_2 would then be residue 51, since no other cyanogen bromide fragment was found to be covalently associated with the NH_2 -

terminal fragment which contained Cys 10. Another potential site for a disulfide could involve Cys residues 266 and 273 of A_2 , but this possibility was not studied. Whether intramolecular cysteines occur in other polypeptides is not known. If present, they must be contained within CNBr fragments since no other mobility shifts were noted when reduced and non-reduced digests were compared electrophoretically.

Each of the major glycinin subunits is synthesized as a precursor polypeptide (8) containing the acidic followed by the basic component (9). A short linker of only a few amino acids is removed during a processing step to yield the individual acidic and basic polypeptides. This processing occurs after the precursor has been sequestered in the protein body (10, 11). Presumably a specific folding of the precursor polypeptide during or shortly after its synthesis facilitates formation of the single disulfide bond which continues to link the two components after the peptide linkage between them is removed. The specificity of the disulfide linkage is emphasized by the fact that in A_{1a} , A_2 , and A_3 cystine residues only 13, 13, and 10 residues away, respectively, from the disulfide were not involved.

One might expect that a significant change in structure would prevent or alter the position of disulfide bond formation. This raises the question of how A_5 , which is only about one-third the size of the other acidic polypeptides, is able to form the secondary and tertiary structure necessary to allow disulfide formation at the same position as in the other subunits. The interchain half-cystine is located about 90 residues from the NH_2 terminus of each acidic polypeptide. However, in A_5 this is only a few residues from the COOH terminus. A likely explanation is that the A_5B_3 subunit is

synthesized as a full sized precursor, but following interchain disulfide formation, the COOH-terminal two-thirds of A₅ is cleaved by a proteolytic enzyme. If this were the case, one might expect to find a protein having an NH₂-terminal sequence which was homologous to the sequence of the other acidic polypeptides at about 90–100 residues from their NH₂ terminus. Interestingly the A₄ NH₂-terminal sequence was approximately 60% homologous to a portion of A₂ which began with residue 100 of A₂ (4). A₄ was the only major acidic glycinin polypeptide not covalently linked to a basic one (1) and a possible explanation for the origin of A₄ would be that it was an internal cleavage product from a precursor containing A₅ at its NH₂ terminus and B₃ at its COOH terminus. The precursor would have to have $M_r \approx 67,000$ to account for all of the sequence contained in A₄, A₅, and B₃. *In vitro* translation products having relative mobilities similar to this have been detected from glycinin mRNAs (8, 9).

One anticipates that A₄, A₅, and B₃ would be present in equimolar amount if they originated from a common precursor. Although A₅ was purified in low amount relative to the other polypeptides (2), its low molecular weight and the presence of 2-mercaptoethanol during its purification may have resulted in substantial loss of this polypeptide. When reducing agents were excluded during subunit purification the recovery of A₅ appeared to be consistent with recovery of A₄ and B₃ (1). The fact that in cultivar Raiden A₄, A₅, and B₃ are simultaneously absent (12) and this coordinate loss is due to a recessive allele at a single genetic locus gives additional support to the hypothesis that A₄, A₅, and B₃ originated from a common primary translation product.³

Although the precise origin of A₄ is uncertain at this time, lack of a disulfide bond would not prevent the A₄ polypeptide from remaining associated with other polypeptides as a subunit of the glycinin molecule as suggested by Iyengar and Ravenstein (13). Noncovalent forces maintain the structure of the glycinin molecule (14) and are undoubtedly involved in

maintaining the structure between each acidic and basic polypeptide in subunits.

Some discrepancies exist between Iyengar and Ravenstein's (13) results and those reported here. These workers reported only three disulfide-linked acidic-basic pairs, whereas we previously found five (1). Furthermore, they claim that A₃ is linked to a polypeptide which they call B₃ and report no evidence for a B₄ polypeptide. This led them to conclude that A₄ is not associated with a basic polypeptide, but is a "free" subunit of glycinin. Since they used an acid-urea gel as the only means to identify the basic polypeptides, and since B₃ and B₄ are not well resolved by this system (2), it is likely that the band which they label B₃ also contains B₄. Thus, we maintain that A₃ is linked by a disulfide bond to B₄, and that B₃ is noncovalently associated with A₄ and covalently linked to A₅.

REFERENCES

1. Staswick, P. E., Hermodson, M. A., and Nielsen, N. C. (1981) *J. Biol. Chem.* **256**, 8752–8755
2. Moreira, M. A., Hermodson, M. A., Larkins, B. A., and Nielsen, N. C. (1979) *J. Biol. Chem.* **254**, 9921–9926
3. Moreira, M. A., Hermodson, M. A., Larkins, B. A., and Nielsen, N. C. (1981) *Arch. Biochem. Biophys.* **210**, 633–642
4. Staswick, P. E., Hermodson, M. A., and Nielsen, N. C. (1984) *J. Biol. Chem.* **259**, 13424–13430
5. Marco, Y. A., Thanh, V. H., Tumer, N. E., Scallan, B. J., and Nielsen, N. C. (1984) *J. Biol. Chem.* **259**, 13436–13441
6. Larkins, B., and Hurkman, W. S. (1978) *Plant Physiol.* **62**, 256–263
7. Laemmli, U. K. (1970) *Nature (Lond.)* **227**, 680–685
8. Tumer, N. E., Thanh, V. H., and Nielsen, N. C. (1981) *J. Biol. Chem.* **256**, 8756–8760
9. Ereken-Tumer, N., Richter, J. D., and Nielsen, N. C. (1982) *J. Biol. Chem.* **257**, 4016–4018
10. Chrispeels, M. J., Higgins, T. J. V., Craig, S., and Spencer, D. (1982) *J. Cell Biol.* **93**, 5–14
11. Chrispeels, M. J., Higgins, T. J. V., Craig, S., and Spencer, D. (1982) *J. Cell Biol.* **93**, 306–313
12. Staswick, P. E., and Nielsen, N. C. (1982) *Arch. Biochem. Biophys.* **233**, 1–8
13. Iyengar, R. B., and Ravenstein, P. (1981) *Cereal Chem.* **58**, 325–330
14. Derbyshire, E., Wright, D. J., and Boulter, D. (1976) *Phytochemistry (Oxf.)* **15**, 3–24

³ K. Kitamura, C. S. Davies, and N. C. Nielsen, manuscript in preparation.