Chapter 27

CHEMISTRY AND BIOLOGY OF THE COTTONSEED GLOBULINS

Julius W. Dieckert
Texas A&M University
College Station, Texas
and
Robert W. Wallace
University of Alabama Medical School
Birmingham, Alabama
and
Marilyne C. Dieckert
Texas A&M Research Foundation
College Station, Texas

INTRODUCTION

The history of the cottonseed globulins is long, but the progress in understanding their chemistry and biology has been slow. As late as 1963 Altschul (1964) pointed to the fact that up to that time supposedly pure protein fractions were heterogeneous by modern analytical criteria. A review of the earlier work on the cottonseed proteins is provided by Altschul *et al.* (1958). Rossi-Fanelli *et al.* (1964) seem to be the first to isolate a monodisperse globulin from cottonseed. They called it "acalin A". Rossi-Fanelli (1968) reported the isolation of a second globulin from cottonseed which he named "acalin B". Berardi *et al.* (1969) described a two-step process that separated the cottonseed proteins into an albumin (Isolate I) and a globulin (Isolate II) fraction. We have established (Wallace, 1976) that Isolate II is a mixture containing mostly acalin A and acalin B.

It seems clear that most of the globulins of the mature cottonseed are localized in the aleurone grains (Hensarling *et al.*, 1969; Martinez *et al.*, 1970). Our ultrastructural analysis indicates that cottonseed aleurins are synthesized and sequestered in a system of endoplasmic reticulum, dictyosomes and vacuoles (Dieckert and Dieckert, 1976a). The process seems common to phylogenetically diverse seed plants including peanuts, shepherd's purse (Dieckert and Dieckert, 1976a) and coconut solid endosperm (Dieckert and Dieckert, unpublished data).

CHEMISTRY OF ACALIN A AND ACALIN B

As mentioned before, acalin A and acalin B are the principal aleurins of cottonseed (Rossi-Fanelli *et al.*, 1964, 1968). We have determined the molecular weight of acalin A by sedimentation equilibrium (Wallace, 1976). Native acalin A in .03 M NaCl, pH 7.0 (NaOH) has a molecular weight of about 119,000 daltons. The same protein in 0.3 M NaCl, 0.1 sodium phosphate, pH 7.0, has a molecular weight of about 198,000 daltons. The dimerization seems to be caused by the presence of phosphate ion and not by a difference in ionic strength. Apparently, Rossi-Fanelli *et al.* (1964) had acalin A in the dimer form, since they report a molecular weight of about 180,000. They also report a value $S^{o}_{20,w}$ of 9.2 for acalin A. A well-behaved spherical protein of 188,300 molecular weight would exhibit such a value. Analysis of native acalin A by ORD shows only about 1.3 percent α-helix (Wallace, 1976). Acalin A appears to be a glycoprotein because it precipitates in the presence of concanavalin A and is bound by concanavalin A attached to sepharose (Wallace, 1976). The bound acalin A is released from the adsorbent by solutions containing α-methyl-D-mannoside. Finally, the major subunit of acalin A gives a positive Schiff test on separation by SDS gel electrophoresis.

Acalin B was less completely analyzed than acalin A. However, analysis by ORD shows it to contain only about 3.6 percent α-helix (Wallace, 1976). The principal subunits of acalin B do not seem to be glycoproteins. However, the best preparations obtained so far contain small quantities of acalin A, a glycoprotein. Both acalin A and acalin B contain several subunits when denatured with SDS. The principal subunits of acalin A and acalin B were isolated from the best available preparations of the parent proteins by the procedure of Wallace *et al.* (1974) and Wallace and Dieckert (1976). The subunit composition of acalin A is given in Table 1. Some of the subunits are present as disulfide-bridged pairs of polypeptide chains. The principal one has a molecular weight of about 98,000 daltons and is composed of two subunits of about 49,000 daltons molecular weight which may or may not be identical. A second, less abundant subunit has a molecular weight of about 80,000 daltons which upon reduction with β-mercaptoethanol yields two subunits of 49,000 and 31,000 daltons. A very minor component exhibits a molecular weight of 77,000 daltons. On reduction subunits of 49,000 and 28,000 are obtained. The most abundant subunit of acalin A has a molecular weight of about 53,500. This subunit is not present as a disulfide-bridged pair and shows no sign of polymerization. Three other subunits with similar properties are of molecular weights 46,000, 31,000 and 18,000 daltons, respectively. These six subunits account for most of the mass of acalin A.

Two disulfide-bridged subunits were isolated from acalin B (Table 2). One of these, P_{35}, has a molecular weight of approximately 35,000 and consists of two subunits of about 20,000 and 15,000 daltons, respectively. The second, P_{42}, has a molecular weight of approximately 42,000 and consists of two disulfide-bridged

Table 1. Molecular weight of the subunits of acalin A of cotton

Protein	State[1]	Mol. wt.[2]	A-S-S-B
P_{98}	not red.	98,000	P_{49}-S-S-P_{49}
P_{49}	red.[3]	49,000	
P_{80}	not red.	80,000	P_{49}-S-S-P_{31}
P_{49}	red.	49,000	
P_{31}	red.	31,000	
P_{77}	not red.	77,000	P_{49}-S-S-P_{28}
P_{49}	red.	49,000	
P_{28}	red.	28,000	
P_{54}	not. red.	53,500	NO
P_{46}	red.	46,000	NO
P_{31}	red.	31,000	NO
P_{18}	red.	18,000	NO

[1]All subunits were denatured with Sodium Dodecyl Sulfate (SDS).
[2]Molecular weights were determined by SDS-page.
[3]Subunits reduced with β-mercaptoethanol.

Table 2. Molecular weight of the subunits of acalin B of cotton

Protein	State[1]	Mol. wt.[2]	A-S-S-B
P_{42}	not red.[3]	42,000	P_{23}-S-S-P_{20}
P_{23}	red.	22,500	
P_{20}	red.	20,000	
P_{35}	not red.	35,000	P_{15}-S-S-P_{20}
P_{15}	red.	15,000	
P_{20}	red.	20,000	

[1]All subunits were denatured with SDS.
[2]Molecular weights were determined by SDS-page.
[3]Subunits were reduced with β -mercaptoethanol.

subunits of 20,000 and 22,000 daltons, respectively. Acalin B contains several other components, probably including acalin A contaminants. The individual reduced subunits of P_{35} and P_{42} have not yet been isolated. The amino acid composition of P_{35} and P_{42} was determined by the same procedures as mentioned below for the acalin A subunits.

POSSIBLE HOMOLOGIES

The amino acid composition of P_{98}, P_{54}, P_{31} and P_{18} of acalin A was determined after hydrolysis with methanesulfonic acid containing 3-(2-aminoethyl) indole (Liu and Chang, 1971; Simpson *et al.*, 1976). The hydrolysis time was varied in order to estimate the labile amino acids and those released with difficulty. Half-cystine was determined as cysteic acid and methionine as methionine sulfone after preformic acid oxidation and hydrolysis with 4 N HCl (Hirs, 1967). Over 93 percent of the dry weight of each subunit was accounted for by the weight of the amino acid residues recovered.

A comparison of the amino acid compositions was made using the Metzger Difference Index (Metzger *et al.*, 1968) as discussed by Dieckert and Dieckert (1976b, 1978). The pairwise comparison of the acalin A subunits is given in Table 3. All of the values of the difference index were less than 9 where a number less than 15 is considered presumptive evidence of genetic homology. The acalin A subunits appear to be genetically related but variable in molecular weight and amino acid composition. We think that several non-allelic structural genes code for the subunits. P_{46} and P_{49} have very similar amino acid compositions; however, P_{46} does not seem to form disulfide-bridged pairs, whereas P_{49} does. The question of the relationship between P_{18} and P_{31} remains open. These proteins may be the products of separate genes or cleavage products of one of the larger subunits.In either case the low values of the difference index for protein pairs involving P_{18} or P_{31} and the other subunits (Table 3) suggest the larger ones may have evolved

Table 3. Metzger indices for the subunits of acalin A[1].

Subunit[2]	P_{18}	P_{31}	P_{46}	P_{49}
P_{54}	5.5	5.2	6.8	7.0
P_{49}	8.2	7.2	2.4	
P_{46}	8.7	7.8		
P_{31}	6.2			

[1]Acalin A from *G. hirsutum* Acala (Wallace, 1976).
[2]The subscript denotes MW x 10^{-3}.

from a much shorter common ancestor by gene duplication and fusion. Definitive answers to these questions should be forthcoming as sequence data on the individual proteins becomes available. (Also see Chapter 28).

A comparison of the amino acid compositions of P_{35} and P_{42} (acalin B) by the Metzger criterion suggests the two proteins are homologous (Table 4). P_{35} cannot be a simple cleavage product of P_{42}, since P_{35} contains more lysine and histidine residues than P_{42}. P_{35} or P_{42} cannot be simple cleavage products of P_{46}, P_{49} or P_{54} of acalin A because there are too many tryptophan and methionine residues. The

Table 4. Comparison of P_{35} of acalin B with legumin-type proteins

Protein		Source	D.I.
P_{22} P_{20}	(42,000)	Acalin B, *G. hirsutum* Acala[1]	8.07
A B	(61,900)	Legumin, *V. sativa*[2]	9.78
A C	(56,300)	Legumin, *V. sativa*[2]	7.59
α β	(58,600)	Legumin, *V. faba*[3]	9.33
P_{25} P_{14}	(67,000)	Arachin, *A. hypogaea*[4]	11.03
Glycinin	(59,600)	*G. max*[5]	8.90
P_{19} P_{30}	(49,000)	*C. nucifera*[6]	9.51

Raw data from: [1]Wallace (1976); [2]Weintraub and Tuen (1971); [3]Wright and Boulter (1974); [4]Yu (1977); [5]Catsimpoolas *et al.* (1971); and [6]Wallace and Dieckert (1976).

preliminary indication is that at least two structural genes code for these proteins and that they are different from the genes coding for the acalin A subunits.

The acalin A and acalin B polypeptides seem genetically related to the reserve proteins of other seeds. The reserve aleurins from phylogenetically diverse species of plants are compositionally and structurally similar. Dieckert and Dieckert (1976b, 1978) adduced evidence that the aleurins are variable but genetically related. Two homology classes were suggested. One, the "vicilin" class, is typified by the vicilins of *Vicia*, α-conarachin of peanuts and concocosin of coconut endosperm. A second class, the "legumins", includes the legumins of *Vicia*, arachin of peanuts, glycinin of soybeans, cocosin of coconut endosperm and edestin of hempseed. The vicilins (Table 5) are usually glycoprotein dimers of low

Table 5. Properties of some proteins of the vicilin type.

Protein	Molecular weight		Glycoprotein	% α-helix
	Parent	SDS subunits		
α-Conarachin: *A. hypogaea*	142,000 (7.85)[1]	72,000[2]	YES[2]	12[5]
Vicilin: *P. aureus*	N.D. (8.05)[4]	50,000[4]	YES[4]	N.D.
Concocosin: *C. nucifera*	110,000 (7.65)[5]	56,000[5]	YES[6]	30[5]
Acalin A: *G. hirsutum*	113,000[7] (9.15) (198,000)	53,500[7] 49,000 46,000	YES[6]	1.3[7]

[1]Dechary *et al.* (1961); [2]Yu (1977); [3]Jacks *et al.* (1973); [4]Ericson and Chrispeels (1973); [5]Khaund (1971); [6]Wallace and Dieckert (unpublished); and [7]Wallace (1976).

N.D. means not determined.

Table 6. Properties of some proteins of the legumin type.

Protein	Molecular weight: Parent	Molecular weight: SDS subunits	α-Helix	Pleated sheet	Unordered	N-terminal residue	$(A\text{-}S\text{-}S\text{-}B)_n$
Legumin: *V. faba*	320,000[1]	21,600 av. (3)[2] 37,000 (2)	N.D.	N.D.	N.D.	Gly[2] Leu, Thr	$(B_{22}\text{-}S\text{-}S\text{-}a_{37})_6$[2]
Glycinin: *G. max*	363,000[3]	22,300 av.[4] 37,000 (3)[5] 45,000	9[6]	33.3	57.7	Gly Phe, Leu (Ile)[5] Leu (Ile)	$(b_{22}\text{-}S\text{-}S\text{-}A_{1,37})_6$[7]
Arachin: *A. hypogaea*	330,000[8]	23,600[9] 43,100	14.6[10]	27.0	58.4	Gly[11] Ile	$(P_{24}\text{-}S\text{-}S\text{-}P_{43})_6$[11]
Edestin: *C. sativa*	309,000[12]	23,000[13] 27,600	6.3[6]	28.1	65.6	Gly[13] Blocked	$(P_{23}\text{-}S\text{-}S\text{-}P_{28})_6$[13]
Cocosin: *C. nucifera*	208,000[14]	19,000[15] 30,000	N.D.	N.D.	N.D.	N.D.	$(P_{19}\text{-}S\text{-}S\text{-}P_{30})_4$[16]

Data drawn from following sources: [1]Bailey and Boulter (1970); [2]Wright and Boulter (1974); [3]Wolf and Briggs (1959); [4]Catsimpoolas *et al.* (1971); [5]Kitamura and Shibasaki (1975a); [6]Jacks *et al.* (1973); [7]Dieckert *et al.*(unpublished); [8]Johnson and Shooter (1950); [9]Singh and Dieckert (1973); [10]Jacks *et al.* (1975), [11]Yu (1977). [12]Svedberg and Pedersen (1940), [13]Dlouhá *et al.* (1964); [14]Sjogren and Spychalski (1930); [15]Wallace and Dieckert (1976); and, [16]Wallace (1976).

N.D. means not determined.

α-helical and high unordered content. The amino acid compositions are similar. Generally, the legumins (Table 6) are hexamers of disulfide-bridged pairs of subunits and usually are not glycoproteins. One member of the pair is an "acidic" and the other a "basic" subunit. As a group, the legumins are conformationally similar with low a-helix, considerable pleated sheet and large unordered conformational components. Compositional similarities between the vicilin and legumin-type proteins suggest that they, too, derive from a common genetic ancestor.

The acalin A proteins probably belong to the vicilin class. They seem to be glycoproteins, their amino acid compositions are similar to other vicilin-type proteins (Table 7) and they appear to have similar three-dimensional structures.

Table 7. Metzger indices for aleurins similar to vicilin from *Phaseolus aureus*.

Subunit[1]	V_{50}[2]	P_{72}[3]	P_{49}[4]
P_{54}[4]	12.3	8.8	7.0
P_{49}	11.0	9.7	
P_{72}	12.9		

[1]Subscript of letter designation is MW x 10^{-3}. Raw data from: [2]Ericson and Chrispeels (1973) (vicilin, *P. aureus*); [3]Yu (1977) (-conarachin, *A. hypogaea*); and [4]Wallace (1976) (two subunits of acalin A from *G. hirsutum*).

Table 8. Comparison of P_{42} of acalin B with legumin-type proteins.

Protein	mol. wt.	Source	D.I.
P_{15} P_{20}	(35,000)	Acalin B, *G. hirsutum* Acala[1]	8.07
A B	(61,900)	Legumin, *V. sativa*[2]	9.39
A C	(56,300)	Legumin, *V. sativa*[2]	10.46
α B	(58,600)	Legumin, *V. faba*[3]	9.32
P_{23} P_{44}	(67,000)	Arachin, *A. hypogaea*[4]	7.34
Glycinin	(59,600)	*G. max*[5]	9.62
P_{19} P_{30}	(49,000)	*C. nucifera*[6]	11.20

Raw data from: [1]Wallace (1976), [2]Weintraub and Tuen (1971); [3]Wright and Boulter (1974), [4]Yu (1977); [5]Catsimpoolas *et al.* (1971); and [6]Wallace and Dieckert (1976).

The acalin B subunits P_{35} and P_{42} seem to belong to the legumin homology class. The non-reduced subunits are composed of two subunits linked by disulfide bridges, and the amino acid composition of each is similar to other legumin-type homologues (Tables 4 and 8). Less is known about the 3-dimensional structure of acalin B proteins than other legumins. However, the available data suggest that native acalin B is structurally similar to the other legumins. We find by ORD studies that acalin B has only about 3.6 percent α-helix. Rossi-Fanelli (1968) observed a sedimentation coefficient of about 11s. This is the value expected for a well-behaved spherical protein of ~246,000 molecular weight. A mixed hex-

amer of equal molar proportions of P_{35} and P_{42} would have a molecular weight of 231,000. These indicators of 3-dimensional structure are similar to those observed for other legumin-type proteins. Acalin A and acalin B may be homologous. They have similar amino acid compositions with differences indices less than 15.

A PROVISIONAL MODEL

From a consideration of the total data there emerges a useful and plausible model for the chemistry and biology of the reserve aleurins of the seed plants. According to the model the reserve proteins are synthesized and sequestered in a system of endoplasmic reticulum, dictyosomes and vacuoles. The aleurins function primarily as a store of organic nitrogen for the seedling. Perhaps structural genes for extracellular proteins were adapted to synthesize the reserve proteins. Apparently there were only a few potential structural genes for the aleurins. Aside from this limitation the requirements for an acceptable reserve protein are minimal. Some possible constraints are: (1) capable of mass production in a short period of time; (2) segregatable from the cytoplasm; (3) high particle weight to permit high tissue concentration without osmotic stress to the cells; and (4) compatible with the germination nitrogen metabolism of the species. Under the model there are many aleurin variants but only a few homology classes. The allowed variability for functionally viable reserve proteins suggests that species may differ with respect to which homology classes are represented. There may be species with vicilins only, legumins only, or with varying proportions of each. Examples of this type of variability may be discerned in Leguminosae.

SUMMARY

Acalin A and acalin B are the principal globulins of the cottonseed. Native acalin A assumes molecular weights of 119,000 or 198,000 depending on the solvent. ORD shows acalin A to contain only about 1.3 percent α-helix. Acalin A appears to be a glycoprotein with several polymorphic forms of subunits. The principal ones have reduced molecular weights of approximately 53,500, 49,000 and 46,000. Comparison of amino acid compositions and other parameters suggest that the subunits are homologous with each other and with the vicilin-type protein of other seed plants. Acalin B is less well characterized than acalin A. It appears to have a native molecular weight of 240,000 to 250,000, 3.6 percent a-helix, no sugars, and consists of hexamers of two kinds of disulfide-bridged pairs of subunits. The molecular weights of the individual pairs are approximately 35,000 and 42,000. Comparison of various properties suggests that these proteins are homologous to the legumin-type protein of other seed plants. The acalin A and acalin B proteins do not seem to be in a precursor-product relationship to each other. A model for the biology and chemistry of the seed globulins is discussed.

ACKNOWLEDGEMENTS

This work was supported by Cotton Incorporated Cooperative Agreement No. 75-682, Texas Agricultural Experiment Station Project H-1306 and Robert A. Welch Foundation Grant No. A-235.