Nishina, Y., Kitagawa, T., Shiga, K., Watari, H., & Yamano, T. (1980a) *J. Biochem. (Tokyo) 87*, 831–839.
Nishina, Y., Shiga, K., Horiike, K., Tojo, H., Kasai, S., Matsui, K., Watari, H., & Yamano, T. (1980b) *J. Biochem. (Tokyo) 88*, 411–416.
Ondrias, M. R., Findsen, E. W., Leroi, G. E., & Babcock, G. T. (1980) *Biochemistry 19*, 1723–1780.
Placzek, G. (1934) in *Rayleigh-Streuung und Raman Effekt, Handbuch der Radiologie*, Heft VI, Teil 2, Akademische Verlagsgesellschaft, Leipzig.
Salmeen, I., Rimai, L., & Babcock, G. T. (1978) *Biochemistry 17*, 800–806.
Spiro, T. G. (1975) *Biochim. Biophys. Acta 416*, 169–187.
Spiro, T. G., & Strekas, T. C. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2622–2626.
Spiro, T. G., & Strekas, T. C. (1974) *J. Am. Chem. Soc. 96*, 338–345.
Spiro, T. G., & Gaber, B. P. (1977) *Annu. Rev. Biochem. 46*, 553–572.
Strekas, T. C. (1976) *Biochim. Biophys. Acta 446*, 179–191.
Vorkink, W. P., & Cusanovich, M. A. (1974) *Photochem. Photobiol. 19*, 205–215.
Yamamoto, T., Palmer, G., Gill, D., Salmeen, I. T., & Rimai, L. (1973) *J. Biol. Chem. 248*, 5211–5213.
Yong, F. C., & King, T. E. (1970) *J. Biol. Chem. 245*, 1331–1335.

Amino Acid Sequence of Ragweed Allergen Ra3†

David G. Klapper,* L. Goodfriend, and J. D. Capra

ABSTRACT: The complete amino acid sequence of ragweed pollen allergen Ra3 has been determined. The molecule consists of 101 amino acid residues and to date is the only allergen isolated from *Ambrosia elatior* (short ragweed) which contains carbohydrate. This particular preparation of allergen has a single, unique amino acid sequence, but there is evidence suggesting that, like Ra5, Ra3 isolated from pollen collected in diverse geographical areas shows amino acid sequence variation. The complete amino acid sequence was derived by utilizing only 10–12 mg of material (approximately 1 μM) as a result of recent technical innovations such as DEAE–glass ion exchangers and Polybrene as a useful sequencing aid. This is the second pollen allergen active in man which has been sequenced, and information resulting from these data should be useful in dissecting the molecular mechanisms involved in atopic allergy.

Sexual reproduction in plants depends in large part upon airborne distribution of pollen from one plant to another. Inhalation of these pollens leads to sensitization of certain individuals such that they develop symptoms of allergic disease (Wodehouse, 1971). Probably the best studied common weed pollen is that from short ragweed (*Ambrosia elatior*). Early studies by King et al. (1964) provided evidence that antigen E was the major allergen of ragweed pollen, but more recently other active allergens have been isolated [King et al., 1967; Underdown & Goodfriend, 1969; Griffiths & Brunet, 1971; Lichtenstein et al., 1973; Lapkoff & Goodfriend, 1974; reviewed by King (1976)]. Although some of these other ragweed allergens are considerably smaller than antigen E and present in only minute quantities, they are quite active and do not appear to cross-react antigenically or allergenically with antigen E (Lichtenstein et al., 1973).

Since ragweed allergens have been implicated in sensitizing certain individuals and eliciting an IgE immune response, it seems important to attempt to understand if ragweed allergens themselves are unique proteins, whether the genetic endowment of a sensitized individual is responsible, or perhaps some combination of these and other factors leads to an allergic state.

The first ragweed allergen to have its complete amino acid sequence determined was Ra5, a 5000 molecular weight protein containing no detectable carbohydrate (Mole et al., 1975). The structure of Ra5 is unusual, having eight cysteines among its 45 amino acids and, at the COOH terminus, three of the last five residues are lysine. The functional relationship of the high cysteine content and very basic "tail" is presently under study.

In addition, hypersensitivity to Ra5 is associated with HLA-B7 histocompatibility antigens (Marsh et al., 1973, 1975). This might suggest that hypersensitivity is simply a "responder" phenomenon and is not a property of the specific allergen involved.

To further investigate this problem, we have determined the primary amino acid sequence of another small ragweed allergen, Ra3 (Underdown & Goodfriend, 1969). This molecule consists of a single polypeptide chain of approximately 100 amino acids and contains some 8% carbohydrate (Goodfriend et al., 1980). An association has been described between sensitivity to Ra3 and the HLA-A2 phenotype (Marsh et al., 1977, 1979).

Materials and Methods

Allergen Ra3. Purified Ra3 was obtained from the Research Resources Branch of the National Institutes of Health. This allergen preparation was extracted from pollen collected by Greer Laboratories (Lanier, NC) and was purified essentially by following the procedures described by Underdown & Goodfriend (1969). For convenience, this preparation will be referred to as NIH Ra3.

Enzyme Digests. Tryptic and chymotryptic digests were performed on extensively reduced (7 M guanidine hydrochloride) and ^{3}H-alkylated Ra3. Dithiothreitol (0.01 M) was added to the protein solution buffered at pH 8.6 with 0.1 M Tris-HCl, and reduction of disulfide bonds proceeded at 37 °C for 30 min. This solution was returned to room temperature and trace labeled in the dark for 15 min with 5×10^{-7}

† From the Department of Bacteriology and Immunology, University of North Carolina, School of Medicine, Chapel Hill, North Carolina 27514 (D.G.K.), the Department of Experimental Medicine, McGill University, Montreal, Canada (L.G.), and the Department of Microbiology, University of Texas Health Science Center at Dallas, Dallas, Texas 75235. *Received February 21, 1980.* This work was supported by National Institutes of Health Grants 1 R01 AE 12796, 1 F32 AI 05018, and 1 R23 AI 14908 and MRC Grant (Canada) MT-2010.

0006-2960/80/0419-5729$01.00/0

M iodo[^{3}H]acetic acid (IAA). At this time, the reaction mixture was brought to 0.02 M IAA and kept dark for 30 additional min. Trace-labeled Ra3 was separated from excess reactants by passage through a G-25 fine Sephadex column equilibrated in 0.2 M NH_4OH. Digests were performed in 1% ammonium bicarbonate solution at 37 °C for 4 h at an enzyme to substrate ratio of 1:100. Citraconic anhydride was used to modify lysine residues prior to tryptic digestion as described previously (Klapper & Capra, 1976).

Chemical cleavage of the two aspartyl–proline (D/P) bonds in this molecule was accomplished in a mixture of 75% formic acid made 7 M in guanidine hydrochloride as recommended by Jauregui-Adell & Marti (1975). The fragments resulting after 12 h at 37 °C were readily separated on a column of Sephadex G-75 equilibrated in 10% acetic acid.

Isolation of Component Peptides. Proteolytic digests of Ra3 were fractionated by molecular sieving through Sephadex equilibrated in 0.2 M ammonium hydroxide. Effluent flow was followed at 280 nm by an ISCO UA5/type 6 recording monitor. Fractions were pooled after sieving, lyophilized, and subjected to high-voltage paper electrophoresis at pH 6.5. Peptides were eluted by descending chromatography in 0.5 M NH_4OH and analyzed on a Durrum D-500 amino acid analzyer after 6 N HCl hydrolysis in vacuo at 110 °C for 24 h. If necessary, peptides were restreaked on Whatman 3MM paper for electrophoretic purification at pH 2.1.

In addition, DEAE-coated glass beads (Corning Glass Works, Medfield, MA) proved to be an invaluable tool in the isolation of several peptides which did not migrate well on paper. The advantage of this methodology is that recovery of material is in excess of 90% of theoretical. Details of the use of DEAE-coated glass have been described (Klapper & Capra, 1976), but in brief, a pyridine–formate decreasing pH (8.6 to 3) and increasing ionic strength gradient was utilized. The column effluent was monitored by the ISCO UA5 system with a duplicate gradient pumped through the reference side of the photometer to compensate for increasing OD_{280} due to increasing pyridine concentration.

Sequencing Procedure. Automated amino acid sequence analysis was performed with a Beckman 890C sequencer by utilizing the recently updated DMAA program. All chymotryptic and most tryptic peptides were sequenced in the presence of Polybrene (Aldrich, Milwaukee, WI) as described in detail elsewhere (Klapper et al., 1977).

The phenylthiocarbamyl derivatives of amino acids from the sequencer were converted to phenylthiohydantoin (PTH) amino acids at 80 °C for 10 min in 0.2 mL of 1.0 M HCl under a nitrogen atmosphere and extracted into ethyl acetate. The aqueous phase was lyophilized, dissolved in 0.2 mL of 1 M ammonium chloride, pH 8.2, and extracted with ethyl acetate to identify histidine and arginine. Toward the end of this study, a Sequemat P6 autoconverter was coupled to the sequencer to perform this step automatically. Aspartic and glutamic acids are converted to their respective PTH methyl esters and are identified as such. PTH amino acids were identified by gas chromatography, thin-layer chromatography, and high-pressure liquid chromatography. Conversion to the parent amino acid was accomplished by hydrolysis in vacuo in HI at 150 °C for 16 h for amino acid analysis. An aliquot (10%) of each step of any radioactive peptide (labeled with ^{3}H at cysteine residues) was counted in a liquid scintillation counter.

Results

NH_2-Terminal Analysis of Ra3. This NIH Ra3 preparation (0.1 μM) and two other preparations isolated from a different batch of pollen were initially subjected to sequencer analysis for approximately 30 steps. Regardless of source, all three sequences were homogeneous and identical, as seen in Table I.

Table I: Amino-Terminal Sequence of Ra3 Isolates (See Text)

position	GC	high-pressure LC	TLC	AAA	sequence
1	G	G	G	G	G
2		K	K	K	K
3	V	V/P/M		V	V
4		Y	Y	Y	Y
5	L/I[a]		L/I	L	L
6	V	V/P/M		V	V
7	G	G	G	G	G
8	G	G	G	G	G
9	P	V/P/M	P	P	P
10		E	E	E	E
11	L/I	L/W	L/I	L	L
12	G	G	G	G	G
13	G	G	G	G	G
14		I/W	W		W
15		K	K	K	K
16	L/I	L/W	L/I	L	L
17		Q	Q	E	Q
18	S	S	S	SAC[b]	S
19		D	D	D	D
20	P	V/P/M	P	P	P
21		R		R	R
22	A	A	A	SAC	A
23		Y	Y	Y	Y
24	A	A	A	SAC	A
25	L/I	L/W	L/I	L	L
26		L/W	W		W
27	S	S	S	SAC	S
28	A	A	A	SAC	A
29					
30		Q	Q	E	Q

[a] X/Y not separated by this technique. [b] S, A, and C all elute as A after back-hydrolysis.

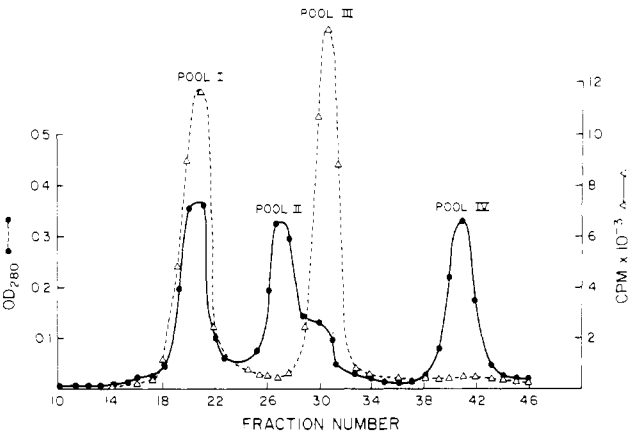

FIGURE 1: Elution profile of extensively reduced, radioalkylated, and citraconylated Ra3 after tryptic digestion and separation through Sephadex G-50 fine equilibrated in 0.2 N NH_4OH.

Major Tryptic Peptides of Ra3. Tryptic peptides from extensively reduced, ^{3}H-alkylated, and citraconylated NIH Ra3 (0.4 μM) were separated through a column of Sephadex G-50 fine equilibrated in 0.2 M NH_4OH. Figure 1 depicts the elution profile monitored at 280 nm. An aliquot (1%) of each tube was counted and ^{3}H activity plotted. The effluent was divided into four pools as shown, and it was noted in three independent digestions that pool III contained approximately 1.2 times the radioactivity of pool I. Pool IV proved to be a pure peptide and was sequenced as the octapeptide AYALW-SAR (Table II, peptide TIV; Table III). Pool II was also relatively pure, contaminated to only a minor extent with

Table II: Major Citraconylated Tryptic Peptides of Ra3

	TIA	TIB	TIC	TII	TIIIA	TIIIB	TIIIB′ [c]	TIV	peptide total	composition (AAA)
Asp	1.1	3.1	2.2	1.1	1.1	1.0			9.6	10
Thr	2.7	3.8							6.5	6.5
Ser	0.9	1.0		1.0				0.9	3.8	5.8
Glu	1.0	4.1	2.1	1.9	1.0	1.0			11.1	10.8
Pro	1.1		1.0	1.8	2.1	1.1	1.0		7.1	6.8
Gly	1.1	1.2		4.6	2.2	1.0	1.1		10.1	9.8
Ala		1.1	2.1			1.0	1.0	2.9	7.1	6.3
Cys[a]	0.7		0.9			0.8	0.7		2.4	2.8
Val		2.7		2.1		1.9	1.8		6.7	5.6
Met										0
Ile	1.1		2.2						3.3	2.6
Leu	2.0	1.1		3.1	1.0			1.0	8.2	6.9
Tyr			0.8	0.8				1.2	2.8	2.0
Phe	1.9	2.8				0.9	1.1		5.6	6.8
His	1.2		1.1						2.3	2.5
Lys	2.0	1.1	1.0	1.9		1.0			7.0	6.8
Arg		1.0	1.0	1.0	1.0	1.0	1.0	1.0	6	4.1
Trp[b]										4

[a] Determined as the carboxymethylated derivative. [b] Not determined for peptides. [c] Not used in determining total (since it is a fragment of IIIB).

Table III: Amino Acid Sequence of Pool IV (See Text)

position	GC	high-pressure LC	TLC	AAA	sequence	yield (nM)
1	A	A	A	SAC	A	60
2		Y	Y	Y	Y	45
3	A	A	A	SAC	A	50
4	L/I	L/W	L/I	L	L	40
5		L/W	W		W	20
6	S	S	S	SAC	S	20
7	A	A	A	SAC	A	40
8		R		R	R	40

Table IV: Amino Acid Sequence of Pool II (See Text)

position	GC	high-pressure LC	TLC	AAA	sequence	yield (nM)
1	G	G	G	G	G	80
2		K	K	K	K	75
3	V	V/P/M		V	V	80
4		Y	Y	Y	Y	75
5	L/I	L/W	L/I	L	L	70
6	V	V/P/M		V	V	75
7	G	G	G	G	G	60
8	G	G	G	G	G	65
9	P	V/P/M	P	P	P	70
10		E	E	E	E	60
11	L/I	L/W	L/I	L	L	40
12	G	G	G	G	G	45
13	G	G	G	G	G	45
14		L/W	W		W	20
15		K	K	K	K	30
16	L/I	L/W	L/I	L	L	35
17		Q	Q	E	Q	25
18	S	S	S	SAC	S	15
19		D	D	D	D	25
20	P	V/P/M	P	P	P	30
21		R		R	R	20

peptides from pool III. Sequence determination (Table IV) and amino acid composition (Table II, peptide TII) proved it to be the NH_2-terminal 1–21 residues. Electrophoresis at pH 6.5 of pool I showed a dark neutral peptide and a major basic peptide. The basic peptide was eventually found to be QQFKTTDVLWFDFTTGEDSVAEVWR (Table II, peptide TIB). It was not possible to determine the entire amino acid sequence of this large peptide, but the sequence was unambiguous for 17 steps. The chymotryptic peptide beginning with TTGED... was sequenced to the end (tryptophan), and the amino acid composition of the large tryptic peptide was in complete agreement with the entire deduced sequence, including the COOH-terminal arginine. This allergen is the only one of the small ragweed proteins to contain carbohydrate. The extremely low, but highly reproducible, yield of aspartic acid at position 12 in this peptide suggests that this is the site of attachment of a carbohydrate moiety. In addition to this major peptide (recovery 45–50% of theoretical), two minor peptides (5–10%) were also recovered. One proved to be QQFK, and the other appeared to be the remainder, i.e., TTDVLWFDFTTGEDSVAEVWR. It is of interest that there was some cleavage of the internal lysine in spite of citraconylation and that the tetrapeptide QQFK appeared in the void volume of the G-50 column. The neutral peptide was electrophoretically further separated at pH 2.1 and yielded two peptides. The composition and the sequence of these peptides varied, depending upon the source of Ra3 pollen. This heterogeneity is of great interest and will form the basis of a separate report. The remainder of this paper will deal with Ra3 obtained from the Research Resources Branch of the NIH. The amino acid sequences of the two remaining peptides from pool I were FTLLTPGSHFICTKDQK (Table II, peptide TIA) and EEAYHACDIKDPIR (Table II, peptide TIC).

Since there theoretically should not have been any cleavage at lysine in this citraconylated molecule, it was at first believed that the peptide FTLLTPGSHFICTKDQK represented the COOH terminus of the molecule. That turned out not to be the case, as will be demonstrated, and leaves the observation that in spite of citraconylation that particular lysine was not modified to any appreciable extent.

Pool III at pH 6.5 gave rise to only one neutral peptide. Subsequent separation at pH 2.1 yielded three peptides, LEPGGPDR (TableII, peptide TIIIA), DQKFVACVPGR (Table II, peptide TIIIB), and FVACVPGR (Table II, peptide TIIIB′). The yield of DQKFVACVPGR was approximately 20% of the yield of FVACVPGR, suggesting incomplete citraconylation of the lysine.

An alternate method of separating citraconylated tryptic peptides was to utilize DEAE-coated glass beads instead of high-voltage electrophoresis. The yields of peptides recovered from paper are variable and range from as little as 35% to only 75%. As an example of the excellent resolution of this resin, Figure 2 depicts the peptide mixture from Ra3 after tryptic

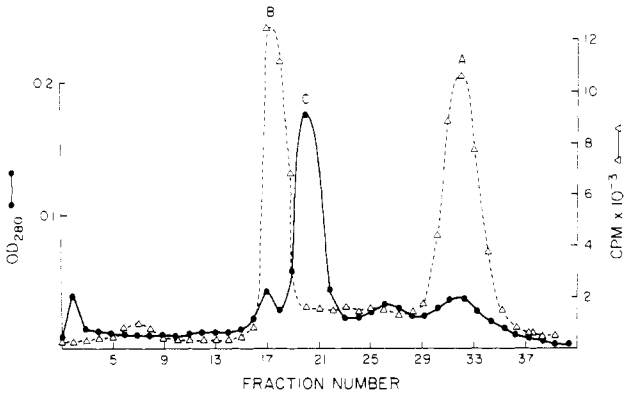

FIGURE 2: Elution profile of the equivalent of peak 1 (Figure 1, tubes 18–22) after separation on DEAE–glass resin. See text and ref 7 for details.

Table V: Chymotryptic Peptides Used To Establish Overlaps

	C1	C2	C3	C4	C5	C6	C7
Asp	1.0		1.0	1.0	3.0		1.0
Thr			1.6	1.8		1.0	0.8
Ser	0.9	0.8		1.0		1.0	
Glu	1.0	2.1		2.0	1.0		1.0
Pro	1.0				3.1	1.0	
Gly				1.2	2.1	1.0	
Ala	1.1	1.1		1.1	1.0		
Cys					1.0		0.9
Val			1.0	2.1			
Ile					2.2		1.0
Leu	1.0		1.1		1.0		
Tyr	1						
Phe		1.1	1.0		1.0	1.0	1.0
His						0.9	
Lys	1.0		1.1		1.0		1.9
Arg	1.0	0.9			1.8		
Trp			+	+			

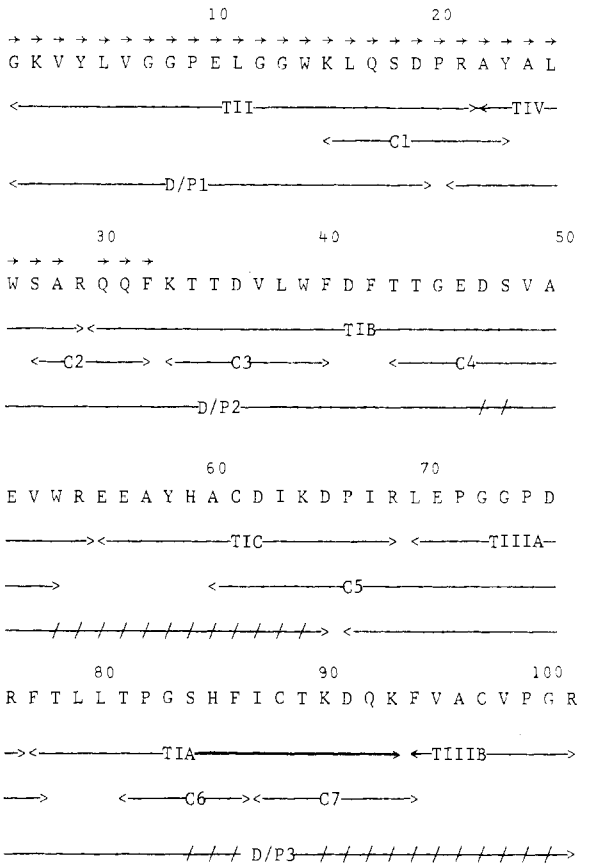

FIGURE 3: Amino acid sequence of NIH Ra3 and the cleavages used to establish overlapping sequences for alignment of component peptides: T, citraconylated tryptic peptide; C, chymotryptic peptide (only useful overlaps shown); D/P, aspartyl–proline cleavage product; /, not identified during sequencer run; →, identified during NH_2 terminal sequencer run.

digestion of citraconylated Ra3. After molecular sieving pool I (Figure 1) was subjected to ion-exchange chromatography on a column (0.6 × 5 cm) of this resin under conditions as described briefly in Materials and Methods. The two radioactive peaks in this preparation were readily separable, as was the large peptide containing two tryptophan residues, and therefore detectable by its absorption properties at 280 nm. Peptide yields after fractionation by this procedure approach 95% of the theoretical value. Peptide A is equivalent to T1A, peptide B is T1C, and peptide C is T1B (Table II).

Major Chymotryptic Peptides of Ra3. A chymotryptic digest of extensively reduced and ^{3}H-alkylated Ra3 (0.4 μM) was treated in much the same manner except that molecular sieving was performed on a column of a mixture of Sephadex G-25 and G-50 fine. The amino acid compositions and sequences of those fragments that helped to establish important sequence overlaps are depicted in Table V and Figure 3, respectively.

Peptide C5 (Table V) is of particular interest since cleavage would be expected to occur after tyrosine-58. In fact, the only high-yield fragment isolated from this region of the molecule had alanine as its amino terminus, suggesting cleavage by chymotrypsin after the histidine.

Another chymotryptic peptide of unusual origin was peptide C6 (Table V). A peptide due to cleavage at phenylalanine-78 and -89 was expected to yield the sequence TLLTPGSHF. Again, contrary to expectation, the only high-yield fragment was TPGSHF, suggesting a major cleavage after the leucine. This turned out to be extraordinarily useful. The serine at position 8 of the tryptic peptide had always been a difficult residue to place because of its extremely low yield. Amino acid composition studies of this peptide suggested a full residue of serine, but since threonine was present at three times the concentration the integration of the serine peak was questioned. However, the amino acid composition of the short peptide TPFSHF is clear (Table VI). There is no question that

Table VI: Amino Acid Composition and Sequence of Chymotryptic Peptide C6

position	high-pressure LC	TLC	AAA	sequence	yield (nM)	amino acid composition	nM	residues
1	T	T	ABU[a]	T	30	T	3.1	1
2	V/P/M	P	P	P	30	S	2.9	1
3	G	G	G	G	25	P	3.0	1
4	S	S	SAC	S	5	G	3.0	1
5	H	H	H	H	20	F	2.8	1
6	I/F	F	F	F	20	H	3.1	1

[a] Aminobutyric acid.

threonine and serine are equimolar. Again, in the sequence of this peptide (Table VI), the yield of serine at step 4 is a fraction of what it should be. This suggests the presence of a carbohydrate moiety at this position.

Disulfide Bonding of Ra3. Addition of iodo[^{3}H]acetic acid to native Ra3 in Tris–NaCl buffer at pH 8.2 did not result in alkylation of this allergen. So, although there are three cysteine residues, some compound (dissociable in the presence of DTE) presumably is blocking the odd sulfhydryl group. In an effort to ascertain which of the cysteines (if any) were disulfide bonded, 100 nM native Ra3 was dialyzed into 1% ammonium bicarbonate and digested with trypsin. The digest was divided equally, and each aliquot was lyophilized. One aliquot was extensively reduced, radioalkylated, and subsequently desalted on a Sephadex G-25 (fine) column equilibrated in 0.2 M NH_4OH. The other aliquot was also subjected to radioalkylation in pH 8.2 Tris–guanidine (but not reduced) and was applied to this molecular sieve. After lyophilization, both aliquots were subjected to high-voltage paper electrophoresis at pH 6.5. The "neutral" region was eluted from the reduced and alkylated preparation and subjected to sequencer analysis. At least five major sequences were present in this radiolabeled eluate, including the peptides EEAY... (glutamic-55) and FTLL... (phenylalanine-78). When a comparable region of the electrophoretic map was eluted from the paper on which the nonreduced digest was placed, neither of these sequences (nor any radioactivity) was present. This map had significant material which failed to migrate from the origin during electrophoresis. Subsequent elution and sequence analysis proved this material to contain both the EEAY... and FTLL... sequences. The peptide FVA... (phenylalanine-97) was found in the neutral fraction in both cases, suggesting that its migration had not been affected by this treatment. The conclusion at this point is that cysteine-61 is disulfide bonded to cysteine-88 and that cysteine-97 is "blocked" by some relatively neutral (in terms of charge at pH 6.5) sulfhydryl compound, but is not involved in intramolecular disulfide bonding. An alternate explanation is that Ra3 actually exists as a dimer, connected by cysteine-97. This possibility is excluded by molecular weight determinations both by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (Adolphson et al., 1978) and by equilibrium sedimentation (L. Goodfriend and M. Roebber, unpublished experiments) which unambiguously yield a molecular weight of approximately 12 000.

Discussion

This study has provided the second amino acid sequence of a ragweed pollen allergen. The major technical innovations which facilitated the determination of the sequence with very limited material available were the use of DEAE-coated glass beads to optimize peptide recovery (Klapper & Capra, 1976) and the use of Polybrene, a quaternary ammonium compound which allows amino acid sequence analysis of either tryptic or chymotryptic peptides up to and including the COOH terminus (Klapper et al., 1977).

The first ragweed allergen which has been sequenced is Ra5, a 45 amino acid residue protein lacking any detectable carbohydrate (Mole et al., 1975). This allergen possessed several structural features which were revealed as a result of primary amino acid sequence analysis. Of 45 total amino acids, 8 residues were cysteines, an extraordinarily high proportion. The COOH terminus of Ra5 is a very basic region of the molecule containing 3 lysines in the last 5 residues. Position two from the NH_2 terminus, in all independent isolates (although all isolates were from pollen collected in the same geographical area), showed two alternative amino acids. That is, approximately one-third of the molecules have leucine in position 2, and two-thirds have valine. This is a relatively conservative alternative amino acid substitution but was the first hint of possible allelic forms of ragweed pollen allergens. As mentioned under Results, Ra3 also demonstrates sequence heterogeneity, although not in the amino terminal 32 residues. Whether these amino acid substitutions are alleles or representative of subspecies of short ragweed remains to be seen. It will be of interest to measure allergic reactions in individuals from one geographical area with Ra3 extracted from pollen obtained from different geographical areas.

The basic COOH terminus of Ra5 has been mentioned previously in this report. With the extensive disulfide bonding in that allergen, it is possible that this highly charged region by virtue of its hydrophilic properties extends out into the medium and may be an allergenic area of this molecule. Ra3 has now been found to possess a COOH-terminal arginine. There are a large number of hydrophobic residues in Ra3, but several areas must represent surface determinants. It is likely the carbohydrate residues attached to aspartic acid-41 and serine-84 are external in nature as is the COOH terminus. The finding that both ragweed allergens studied in such detail have very basic COOH terminal residues may imply that this region of the molecule is the site of attachment to the pollen grain by means of ionic interactions. It is also of interest that attempts to sequence from the carboxyl terminus of Ra3 have not been successful. A 50 nM aliquot of reduced and alkylated Ra3 after up to 12 h of carboxypeptidase Y treatment yielded no cleavage product. Therefore, it may be that the COOH-terminal arginine is modified in some way, i.e., by ethanolamine or another compound covalently coupled to the carboxyl group. This has not yet been investigated, nor has any further study of the carbohydrate been initiated simply due to the lack of appreciable quantities of Ra3.

The long-range goal of these studies is to isolate and characterize those areas of the molecule implicated in allergic disease in man. It has been found that the two aspartic acid–proline doublets in the molecule are amenable to acid cleavage with the generation of three large peptides representing the entire molecule (Figure 3). These peptides are well separated on a column of Sephadex G-75 and will be used to establish which peptides mimic the allergic activity of intact Ra3.

Now that the amino acid sequence of an Ra3 molecule has been established, studies can begin to alter allergenic activity by modification of residues as has been done with antigen E, another ragweed allergen [reviewed by King (1976)]. The advantage of performing these studies with Ra3 is that on the basis of these sequence studies it may be easier to ascertain what modifications are more likely to result in useful alteration of the molecule.

Acknowledgments

We thank T. Wallace, F. E. Page, and S. McQueen for excellent technical assistance. In addition, the patient secretarial help of J. Goldman is greatly appreciated.

References

Adolphson, C., Goodfriend, L., & Gleich, G. J. (1978) *J. Allergy Clin. Immunol. 62*, 197.

Goodfriend, L., Roebber, M., Lundkvist, U., & Choudhury, A. (1980) *J. Allergy Clin. Immunol.* (in press).

Griffiths, B. W., & Brunet, R. (1971) *Can. J. Biochem. 51*, 1275.

Jauregui-Adell, J., & Marti, J. (1975) *Anal. Biochem. 69*, 468.

King, T. P. (1976) *Adv. Immunol. 23*, 77.
King, T. P., Norman, P. S., & Connell, J. T. (1964) *Biochemistry 3*, 458.
King, T. P., Norman, P. S., & Lichtenstein, L. M. (1967) *Biochemistry 6*, 1992.
Klapper, D. G., & Capra, J. D. (1976) *Ann. Immunol. (Paris) 127C*, 261.
Klapper, D. G., Wilde, C. E., & Capra, J. D. (1977) *Anal. Biochem. 85*, 126.
Lapkoff, C. B., & Goodfriend, L. (1974) *Int. Arch. Allergy Appl. Immunol. 46*, 215.
Lichtenstein, L. M., Roebber, M., & Goodfriend, L. (1973) *J. Allergy Clin. Immunol. 51*, 285.
Marsh, D. G., Bias, W. B., Hsu, S. H., & Goodfriend, L. (1973) *Science (Washington, D.C.) 179*, 691.
Marsh, D. G., Bias, W. B., Santilli, J., Schachter, B., & Goodfriend, L. (1975) *Immunochemistry 12*, 539.
Marsh, D. G., Goodfriend, L., & Bias, W. B. (1977) *Immunogenetics (N.Y.) 5*, 217.
Marsh, D. G., Chase, G. A., Freidhoff, L. R., Meyers, D. A., & Bias, W. B. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2903.
Mole, L. J., Goodfriend, L., Lapkoff, C. B., Kehoe, J. M., & Capra, J. D. (1975) *Biochemistry 14*, 1216.
Underdown, B. J., & Goodfriend, L. (1969) *Biochemistry 8*, 980.
Wodehouse, R. P. (1971) *Hayfever Plants*, 2nd ed., Hafner Publishing Co., New York.

Behavior of Glycopolypeptides with Empirical Molecular Weight Estimation Methods. 1. In Sodium Dodecyl Sulfate†

Bonnie Strayer Leach, James F. Collawn, Jr., and Wayne W. Fish*

ABSTRACT: The influence of the presence of oligosaccharide branches was examined with respect to the behavior of glycopolypeptides in empirical molecular weight estimation methods in the presence of sodium dodecyl sulfate ($NaDodSO_4$). This examination was conducted by comparing the gel chromatographic and gel electrophoretic behaviors in the presence of $NaDodSO_4$ of 13 glycopolypeptides of known chemical and physical properties to those of regular polypeptides. Errors in the gel chromatographic molecular weight for glycopolypeptides in $NaDodSO_4$ varied from −22% to +10% and indicated that the hydrodynamic behavior of the glycopolypeptide–$NaDodSO_4$ complex could not be correlated with the amount of carbohydrate in the glycopolypeptide. $NaDodSO_4$ binding measurements on a number of the glycopolypeptides suggest that the polypeptide moiety binds the nominal weight ratio of $NaDodSO_4$, while the carbohydrate portion exhibits little or no $NaDodSO_4$ binding. As has been reported by others, the polyacrylamide gel electrophoretic behavior of glycopolypeptide–$NaDodSO_4$ complexes yielded abnormally high molecular weight estimates. In general, the error of these estimates diminished with decreasing porosity of the gel; however, each glycopolypeptide behaved in a unique fashion. Treatment of the electrophoretic data by any of several empirical means provided no reliable way to correct for the glycopolypeptides' aberrant behavior.

Only recently have we recognized the ubiquitous distribution of glycoproteins throughout biological systems. As a first step in gaining an accurate understanding of the functional importance of the covalently attached carbohydrate chains, the glycoprotein must be accurately characterized chemically and physically. Physical characterization includes, of course, the most fundamental property, molecular weight. Empirical methods for polypeptide chain molecular weight estimation currently enjoy widespread use; however, for the results to be reliable, these empirical methods demand strict adherence to certain requirements (Fish, 1975). The primary of these requirements is that the effective hydrodynamic size of a polypeptide be a unique function of its chain length and, hence, its mass. Gel electrophoresis in the presence of sodium dodecyl sulfate[1] additionally requires that the electrostatic charge on the polypeptide–detergent complex must be proportional to the mass of the polypeptide. These conditions are generally met by *linear polypeptides* complexed in the normal fashion with $NaDodSO_4$ (Reynolds & Tanford, 1970a,b). However, by their very chemical nature, glycopolypeptides as branched-chain polymers cannot meet this primary conformational requirement. In spite of this fact, attempts to estimate the molecular weights of glycopolypeptides by making empirical corrections to $NaDodSO_4$–gel electrophoresis data are frequently made.

Quantitatively, it is not known just how much the structural and chemical incongruity of glycopolypeptides affects their behavior in $NaDodSO_4$ denaturing solvent systems for empirical molecular weight estimations. In an attempt to obtain information relative to this problem, we have compared the gel chromatographic and gel electrophoretic behaviors of a number of glycopolypeptides of known physical and chemical properties to those of regular polypeptides. The results of these comparisons are presented herein.

Materials and Methods

Glycopolypeptides. α_1-Acid glycoprotein was obtained from Miles Laboratories as human glycoprotein, fraction VI (lot II). The preparation exhibited a single electrophoretic zone

† From the Department of Biochemistry, Medical University of South Carolina, Charleston, South Carolina 29403. *Received May 5, 1980.* This research was supported in part by a grant from the Research Corporation.

[1] Abbreviations used: $NaDodSO_4$, sodium dodecyl sulfate; BSA, bovine serum albumin; DNP, dinitrophenyl.

0006-2960/80/0419-5734$01.00/0