Phytochemistry, 1980, Vol. 19, pp. 2113–2115. Pergamon Press Ltd. Printed in England.

N-TERMINAL AMINO ACID SEQUENCE OF *C* HORDEIN

PETER R. SHEWRY, JOHN F. MARCH* and BENJAMIN J. MIFLIN

Rothamsted Experimental Station, Harpenden, Herts. U.K.; *John Innes Research Institute, Norwich, U.K.

(*Received* 11 *February* 1980)

Key Word Index –*Hordeum vulgare*; Gramineae; barley; prolamin; hordein; amino-acid sequence.

Abstract—The *C* hordein (prolamin storage protein) fraction of barley endosperm has been purified and the *N*-terminal sequence of amino acids determined for 30 residues. No sequence was obtained for the *B* hordein fraction because the *N*-terminus was blocked.

INTRODUCTION

The prolamin fraction of cereal endosperms has been the subject of much recent research (see reviews [1, 2]). In particular a number of partial amino acid sequences have been published of different components of the prolamin fraction of wheat (gliadin) [3–5] and maize (zein) [6, 7]. Little information is available on the sequence of the prolamin fraction of barley (hordein). Hordein polypeptides can be classified into two major groups termed *B* and *C*; these groups have been shown to be under the control of two separate, but linked, loci on chromosome 5 termed Hor-2 and Hor-1 respectively [8, 9]. These two fractions also differ in a number of properties particularly in their amino acid composition, their solubility and the number of component polypeptides. This paper reports initial studies on the amino acid sequence of these different groups in which the *N*-terminal sequence of 30 amino acid residues of the *C* hordein fraction has been determined.

RESULTS AND DISCUSSION

The *N*-terminal amino acid sequence ascribed to the *C* hordein preparation is given in Table 1. The recoveries obtained are somewhat lower than might have been expected, being 41 and 29% on two duplicate runs for the first residue, as calculated from regression analysis. The reasons for this may be many-fold: for example, one or more of the four polypeptides known, by two-dimensional electrophoresis to be present in the *C* hordein fraction of Julia, may have been blocked; also the coupling of prolamins during sequencing may be less efficient than with other proteins. The main detection systems used do not deal satisfactorily with arginine, pyridylethylcysteine or serine. However, based on amino acid analysis of the *C* fraction [10] pyridylethylcysteine is unlikely to be encountered as it was only detected in trace amounts, and often some indication of the presence of serine can be obtained; where this occurred a tentative assignment of serine has been made. The identification of the *N*-terminal residue as arginine is based on dansylation experiments [10] and on back-hydrolysis of the phenylthiohydantoin (PTH) amino acid released from the sequencer run with subsequent analysis by cation exchange chromatography. Preliminary studies of total hordein preparations by Bietz (personal communication) have also indicated only arginine as an *N*-terminal amino acid. Regression analysis based on the recoveries of leucine and proline at residues 3, 10, 13, 18 and 28 gave a straight line relationship with a repetitive yield of 95.4%. This indicates that these amino acids are probably the only ones present in this position. Certain other positions appear however to have more than one amino acid present. The yield of proline at position 5 was lower than expected although no other residue could be identified. A similar low recovery of proline was noted at positions 12, 20, 21 and 23; for the latter there was also evidence for the presence of phenylalanine. One reason for the presence of proline at positions 12 and 20 could be a certain amount of premature cleavage of the subsequent residue giving a preview effect and in a duplicate run there was no evidence for proline at position 12.

The frequency of residues occurring in the first 30 positions is consistent with the known overall composition of the proteins in which glutamine and proline predominate [10]. When the results are compared with other known sequences of partially purified prolamins there is no evidence of sequence homology (Table 1). There is also no homology with partial sequences of bulk gliadin fractions from *Triticum dicoccoides*, *Aegilops squarrosa* or *Secale cereale* [4]. This is, perhaps, not too surprising since there is little evidence for similarity in the amino acid composition of *C* hordein [10], α- and γ-gliadin [11] and zein [12] (e.g. proline contents are respectively 30, 13, 16 and 11%, phenylalanine 8.8, 3.6, 4.6 and 5.0%, and *S*-amino acids 0.2, 4.0, 3.0 and 1.3%). However, there is a strong similarity between the amino acid composition of *C* hordeins [10] and certain ω-gliadins (proline 20–30%, phenylalanine 8–10% and *S*-amino acids 0–0.4%) [13] and it will be interesting to compare the sequence homology of these two groups before making any premature conclusions as to sequence divergency between the cereal storage proteins from different species (see *Note added in proof*).

Attempts were also made to sequence the *B* hordeins but with no success. It would appear from these studies and from dansylation and hydrolysis that the *B* polypeptides are blocked in the *N*-terminal positions. So far attempts to unblock the terminus have proved unsuccessful.

Table 1. *N*-terminal amino acid sequence of *C* hordein in comparison with other known prolamin sequences

Sample	Residue number																												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
C Hordein Primary Sequence	R	Q	L	N	P	S	S	Q	E	L	Q	S	P	Q	Q	S	Y	L	Q	P	P	Y	P	Q	N	P	Y	L	P	P
Secondary Residues					?							P				P				F	Q		F		Q		P			T
Recovery (nmol)*	37		37		15							P=5	29					23		P=9	P=10		P=7					11		
α-$_{11}$Gliadin† Primary Sequence	V	R	V	P	V	P	Q	L	Q	P	Q	N	P	S	Q	Q	Q	P	Q	E	Q	V	P	L	V	Q	Z			
γ-$_2$Gliadin‡ Primary Sequence	N	I	G	V	D	P	W	G	Q	V	Q	W	L	P	Q	Q	Q	V	P	Q	L	Z	Q	P						
Zein§ Primary Sequence	T	I	F	P	Q	C	S	Q	A	P	I	A	I	L	L	Q	P	Y	L	P	V	A	V	M	G	V	Q	Z	N	A

The notation of amino acids follows the standard abbreviations. Residues underlined are subject to uncertainty.

* 154 nmol loaded based on 90% dry wt of sample and assuming a MW of 52000.

† Taken from ref. [5], five other α-gliadins, one β-and two γ-gliadins have virtually identical sequences [5, 3].

‡ Also from ref. [5], one other γ-gliadin has a similar sequence.

§ Taken from ref. [6], a similar sequence is reported in ref. [7].

EXPERIMENTAL

The *B* and *C* hordein fractions were derived from dry seed of *Hordeum vulgare* cv Julia by the methods described by Shewry *et al.* [10]. They were of corresponding purity as judged by SDS-PAGE to the preparations reported on in that paper.

N-terminal amino acid sequence analysis was carried out with a Beckman 890C automatic sequencer. Pyridylethylated hordein (*ca* 10 mg) was applied to the sequencer in 0.5 ml 10% (v/v) HOAc. A 1.0 M quadrol buffer system (Beckman programme No. 122974) was used, with double coupling on the first degradation cycle. The resultant thiazolinones were converted to the PTH amino acids [14], and identified using GLC [15] and TLC [16]. The first four residues were also back-hydrolysed with 6 N HCl containing $SnCl_2$ for 4 hr at 150° [17] to their parent amino acids which were then examined on a Beckman 120 amino acid analyser [18].

Acknowledgement—We are pleased to acknowledge the helpful advice of Dr. D. Kasarda.

REFERENCES

1. Miflin, B. J. and Shewry, P. R. (1979) in *Seed Protein Improvement in Cereals and Grain Legumes* Vol. 1, pp. 137–158. IEAE, Vienna.
2. Wall, J. S. (1979) in *Recent Advances in the Biochemistry of Cereals* (Laidman, D. L. and Wyn Jones, R. G., eds.) p. 275. Academic Press, London.
3. Kasarda, D. D., Da Roza, D. A. and Ohms, J. I. (1974) *Biochim. Biophys. Acta.* **351**, 290.
4. Autran, J. C., Lew, E., J.-L., Nimmo, C. C. and Kasarda, D. D. (1979) *Nature* **282**, 527.
5. Bietz, J. A., Huebner, F. R., Sanderson, J. E. and Wall, J. S. (1977) *Cereal Chem.* **54**, 1070.
6. Bietz, J. A., Paulis, J. W. and Wall, J. S. (1979) *Cereal Chem.* **56**, 327.
7. Larkins, B. A., Pedersen, K., Handa, A. K., Hurkman, W. J. and Smith, L. D. (1979) *Proc. Natl. Acad. Sci.* **76**, 6448.
8. Shewry, P. R., Pratt, H. M., Finch, R. A. and Miflin, B. J. (1978) *Heredity* **40**, 463.
9. Doll, H. and Brown, A. D. H. (1979) *Can. J. Genet. Cytol.* (in press).
10. Shewry, P. R., Field, J. M., Kirkman, M. A., Faulks, A. J. and Miflin, B. J. (1980) *J. Exp. Botany* **31**, 393.
11. Charbonnier, L. (1973) *Biochimie* **55**, 1217.
12. Sodek, L. and Wilson, C. M. (1971) *Agric. Food Chem.* **19**, 1144.
13. Charbonnier, L. (1974) *Biochim. Biophys. Acta.* **359**, 142.
14. Edman, P. and Begg, G. (1967) *Eur. J. Biochem.* **1**, 80.
15. Pisano, J. J., Bronzert, T. J. and Brewer, H. B., Jr. (1972) *Analyt. Biochem.* **45**, 43.
16. Kulbe, K. D. (1974) *Analyt. Biochem.* **59**, 564.
17. Mendez, E. and Lai, C. Y. (1975) *Analyt. Biochem.* **68**, 47.
18. Spackman, D. H., Stein, W. H. and Moore, S. (1958) *Analyt. Chem.* **30**, 1190.

Note added in proof: Essentially the same sequence as reported here has now been found in bulk hordeins [Bietz, J. *Cereal chem.* (in press)] and in an ω-gliadin from *Triticum monococcum* [Shewry, P. R., Autran, J. C., Lew, E., Nimmo, C. C. and Kasarda, D. D. *Nature* (in press)].