A conformational study of a glutamine- and proline-rich cereal seed protein, C hordein

Arthur S. TATHAM,*† Alex F. DRAKE† and Peter R. SHEWRY*

*Department of Biochemistry, Rothamsted Experimental Station, Harpenden, Herts. AL5 2JQ, U.K., and

†Department of Chemistry, University of London King's College, Strand, London WC2R 2LS, U.K.

(Received 19 September 1984/Accepted 6 November 1984)

A combination of c.d. spectroscopy and computer prediction is used to show that C hordein has an unusual secondary structure with an absence of α -helix and β -sheet, but the presence of regularly repeated β -turns. This is associated with a repetitive primary structure based mainly on blocks of eight residues. Similar spectral changes occurred when the protein was heated from 6 to 86°C in aq. 70% (v/v) ethanol or dissolved in increasing concentrations (50–100%, v/v) of trifluoroethanol in water. The studies indicated that the conformation is stabilized by strong hydrophobic interactions and by extensive hydrogen-bonding.

C hordein consists of a group of storage proteins that together account for up to 10% of the total proteins of mature barley grain. It has the unusual property, shared by other cereal seed prolamins (Mifflin & Shewry, 1979), of being insoluble in water or dilute salt solutions, but soluble in mixtures of alcohols (ethanol, propan-1-ol, propan-2-ol) and water. This is associated with an unusual amino acid composition, high in glutamine (40 mol/100 mol) and phenylalanine (9 mol/100 mol), low in acidic and basic amino acids (about 5 mol/100 mol in total) and no cysteine (Shewry *et al.*, 1980a).

Although C hordein is a mixture of proteins, these are known to have a high degree of structural homology. Available amino acid sequences indicate that they have a regular structure with short *N*- and *C*-terminal domains and a central core consisting of repeated blocks of five and eight residues. In the present paper we use a combination of structural prediction and c.d. spectroscopy to show that the proteins have a regular, but unusual, secondary structure consisting of repetitive β -turns.

Materials and methods

Materials

A total C hordein fraction was purified from Risø 56, a mutant barley line with an increased

proportion of this group of proteins (Kreis *et al.*, 1983), as described previously (Shewry *et al.*, 1981). Two-dimensional isoelectric focusing–sodium dodecyl sulphate/polyacrylamide-gel electrophoresis showed the presence of about ten major polypeptides in this fraction (Kreis *et al.*, 1983). Analysis has shown almost identical amino acid compositions and *N*-terminal sequences (Shewry *et al.*, 1980a, 1981). The total fraction can therefore be treated as a single protein for physical-chemical analysis.

Reagents used were of analytical grade, and c.d. solvents were of the highest purity available. Protein c.d. solutions were filtered before use (Millex GV filters, 0.22 μ m-pore-size, Millipore Corp.).

Methods

C.d. measurements were made with a Jasco J40CS dicograph. Variable-temperature spectra were obtained by using a heated cell holder and a thermocouple of a Comark electronic thermometer inserted directly into the solution. The results are calculated by using an average residue weight of 105.5, calculated from the amino acid composition (Shewry *et al.*, 1980a); the units are degrees $\cdot$ cm² $\cdot$ mol⁻¹. Absorption spectra of the same solutions were recorded with a Cary 210 spectrophotometer. The concentration of protein was calculated from the $A_{280}^{1\%}$ in aq. 50% (v/v) propan-1-ol of 7.96 (D. D. Kasarda, personal communication).

The secondary structures were predicted from the available amino-acid-sequence data by using

Abbreviation used: TFE, trifluoroethanol.

† To whom correspondence and requests for reprints should be addressed.

two methods, namely those of Chou & Fasman (1978) and Garnier *et al.* (1978). For the Chou & Fasman (1978) analysis, predictions were obtained from the products of N_α and N_β of the residue parameters P_α and P_β (Dufton & Hider, 1977). Search distances of 6 and 5 were used for helical and sheet regions respectively. For the β -turn analysis, Chou & Fasman (1978) calculated the average probability of a turn occurrence (pt) as 0.55×10^{-4} and selected tetrapeptides with $pt > 0.75 \times 10^{-4}$ as probable turns. For the method of Garnier *et al.* (1978) the unweighted prediction results were used.

Results and discussion

Prediction of secondary structure from primary structure

The *N*-terminal amino acid sequences of total C hordein and purified fractions show a high degree of homology (Schmitt & Svendsen, 1980; Shewry *et al.*, 1980a,c, 1981; Shewry & Mifflin, 1983), with only one substitution in the first 30 residues. The sequences show a short *N*-terminal domain of 12 residues followed by three blocks of five residues, with a consensus sequence of Pro/Leu-Gln-Gln-Pro-Tyr (Fig. 1a). The preparation and characterization of two random chymotryptic peptides has been reported (Forde *et al.*, 1985); both have an octapeptide repeat with a consensus sequence of Pro-Gln-Gln-Pro-Phe-Pro-Gln-Gln (Fig. 1b,c). Since phenylalanine can be substituted for tyrosine by a single base change, the octapeptide probably corresponds to the *N*-terminal pentapeptide with the addition of a tripeptide. It is thought that the octapeptide repeat motif extends through most of the protein: the consensus composition of Gln₄Pro₃Phe₁ is close to the proportions of these residues in C hordein (Shewry *et al.*, 1981), whereas nucleotide sequencing of cloned complementary DNA species has shown that the octapeptide repeat extends to six residues from the *C*-terminus (Forde *et al.*, 1985).

Prediction by the method of Chou & Fasman (1978) indicates that the β -turn conformation is the dominant structural feature in the C hordein sequences (Fig. 1). This is confirmed by the method of Garnier *et al.* (1978), which predicts β -turn and random-coil conformations (results not shown).

Repetitive β -turns are predicted for the *N*-terminal pentapeptide repeat (Fig. 1a), the probabilities depending on the tetrapeptide sequence. Two overlapping repetitive β -turns are predicted for the consensus octapeptide repeat (Fig. 1d), sharing a glutamine residue and separated by a glutamine residue.

The absence of predicted α -helical or β -sheet

structures in the repeat sequences can be accounted for by the high content of proline and its occurrence every second, third or fourth residue (Figs. 1a, 1b and 1c). A minimum of five favourable residues are required to form an α -helix and six a β -sheet, proline destabilizing either structure (Chou & Fasman, 1978). Thus neither structure can form. A polyproline type-II structure is also unlikely to form, requiring consecutive proline residues or proline residues at every third position along the polypeptide chain (Isemura *et al.*, 1983; Shibata *et al.*, 1984).

Determination of secondary structure from c.d. spectra

C hordein is sparingly soluble in water and the c.d. spectrum was initially determined in aq. 70% (v/v) ethanol. The far-u.v. spectrum (Fig. 2a) is dominated by transitions associated with the amide groups of the peptide backbone, and despite the relatively high content of phenylalanine in C hordein, aromatic contributions are expected to be too weak to contribute significantly to the c.d. in the 230–180 nm region (Woody, 1978). The most striking feature in this region is the absence of characteristics associated with either an α -helical- or β -sheet-rich conformation (Manavalan & Johnson, 1983), supporting the results of structure prediction. Indeed the spectrum most closely resembles that of a random coil (Greenfield & Fasman, 1969). However, when the spectrum in aqueous ethanol of C hordein is compared with that in 4.0M-urea, a protein denaturant, differences are apparent (Fig. 2a). The spectrum in aq. 70% ethanol shows a shoulder at 230 nm, with a negative peak centred at 203 nm, whereas the shoulder is absent in that in 4.0M-urea. The position of the negative peak could not be determined in this solvent, owing to the absorption of urea at low wavelengths (Fig. 2a). The near-u.v. spectrum in aq. 70% ethanol showed characteristic phenylalanine band structure, with two vibronic components centred at 262 and 268 nm (Strickland, 1974). The phenylalanine contribution in this region is relatively strong. This would not be expected with a free-rotating short-axis phenyl electronic transition, indicating that the phenylalanine side chains adopt a fixed conformation. This in turn is indicative of an ordered protein structure. In 4.0M-urea there is complete loss of phenylalanine band structure in the near-u.v. region. These observations in the far- and near-u.v. regions indicate that the conformation in aq. 70% ethanol is not random.

The effect of heat on the conformation of C hordein in aq. 70% ethanol was studied. Increases in temperature would be expected to result in increasing disorder and decrease the intensity of,

(a) *N*-Terminus

(b) Chymotryptic peptide 1

(c) Chymotryptic peptide 2

(d) Consensus octapeptide repeat

Probabilities of predicted turns

a, 6.42×10^{-4}	e, 1.73×10^{-4}
b, 1.74×10^{-4}	f, 3.87×10^{-4}
c, 1.32×10^{-4}	g, 0.81×10^{-4}
d, 0.82×10^{-4}	h, 0.98×10^{-4}

Fig. 1. Structure prediction on the available sequences of *C hordein*

The horizontal line indicates the position of a predicted β -turn. The full *N*-terminal sequences of the two chymotryptic peptides have not been published previously.

and blue-shift, the negative peak at 203nm (Tanford, 1968). When the temperature was increased from 6 to 86°C the differential absorption increased monotonically, and the position of the peak was red-shifted to 205–206nm (Fig. 3). There was also a concomitant decrease in the intensity of the shoulder centred at 230nm. The set of temperature curves exhibited an isocircular dichroic point at 213nm. This suggests that there

are two defined conformations, one at high and one at low temperature, with a negligible concentration of intermediates. The near-u.v. spectra, however, showed very little change over the same temperature range (spectra not shown) (see Fig. 2b). This can only be explained if the secondary structure is stabilized, at least partially, by hydrophobic interactions, which are favoured at high temperatures (Tanford, 1970). *C hordein* neither

Fig. 2. C.d. of *C hordein* in aq. 70% (v/v) ethanol (—) and 4.0M-urea (---). (a) Far-u.v. spectrum; (b) near-u.v. spectrum. The protein concentration was 1 mg/ml.

precipitated nor formed a turbid solution at 86°C, and the effect of heat on the c.d. spectra was completely reversible.

Increasing the concentration of TFE, a solvent which enhances intramolecular interactions, such

Fig. 3. Far-u.v. c.d. spectra of *C hordein* with increasing temperature.

The solvent was aq. 70% (v/v) ethanol; the protein concentration was 1 mg/ml.

as hydrogen-bonding and electrostatic salt bridges, in peptides and proteins, from 50 to 100% (v/v) in water resulted in similar changes in the far- and near-u.v. regions as those observed on heating. These are an increase in the intensity and a shift in the peak wavelength from 203 to 205–206 nm, an isocircular-dichroic point at 213 nm and a decrease in the intensity of the shoulder at 230 nm. A far-u.v. difference spectrum was obtained by subtracting the spectrum in aq. 50% (v/v) TFE from that in 100% TFE (Fig. 4). A similar spectrum could be obtained by subtracting the 6°C spectrum from the 86°C spectrum in Fig. 3 (spectrum not shown). The difference c.d. spectrum corresponds to a class-B β -turn spectrum according to the classification scheme of Woody (1974), with a minimum around 228 nm and maximum in the region 200–205 nm (Fig. 4). This class-B spectrum is the most common class associated with any β -turn conformation, and its observation is indicative of the β -turn conformation (Smith & Pease, 1980). The increase in β -turn conformation content with both increasing temperature and decreasing hydrophobicity of the

Fig. 4. Far-u.v. difference c.d. spectrum obtained by subtracting the spectrum in aq. 50% (v/v) TFE from that in 100% TFE

solvent would indicate that the β -turn-rich conformation maximizes both hydrophobic interactions and hydrogen-bonding.

General discussion

The conformation of C hordein in 100% TFE may resemble that of the native protein in the protein body of the seed, where the water content is low, and hydrogen-bonding and hydrophobic interactions are favoured. The c.d. spectra closely resemble those of proteins that are rich in β -turns. Synthetic-model elastin polypeptides and the sea-anenome toxin anthopleurin A are both rich in β -turns, and possess c.d. spectra with minima around 202–204 nm and shoulders between 220 and 230 nm (Urry *et al.*, 1974; Ishizaki *et al.*, 1979). From the theoretical β -turn backbone conformations described by Venkatachalam (1968), Woody (1974) calculated four different classes of c.d. spectra. Thus there is no unique c.d. spectrum associated with the β -turn conformation. Experimental evidence for these spectral classes comes from studies on synthetic-model β -turn containing peptides (Bush *et al.*, 1978; Lisowski *et al.*, 1983; Urry *et al.*, 1974). Proteins rich in β -turns may therefore be expected to contain mixtures of turn types, which would be expected to differ from model types owing to the effects of solvent, local environment and constraints on the β -turn within the protein.

This would be reflected in the c.d. spectra. Because C hordein contains no cysteine, the β -turn-rich conformation is probably stabilized by hydrophobic interactions between aromatic residues and by hydrogen-bonding between amino and carboxy groups of glutamine residues with other glutamine residues and possibly the peptide backbone. Increasing the temperature or the hydrophobicity of the solvent would affect the contributions of these non-covalent interactions, resulting in conformational and spectral changes.

In elastin the presence of a β -turn in a polypentapeptide repeat has been suggested to result in a linear helical structure called a β -spiral (Venkatachalam & Urry, 1981). The β -turn in the pentapeptide repeat in the *N*-terminal domain of C hordein could also form such a helical structure, with Pro-Tyr as the central residues of the turn, and with a glutamine residue spacing the turns (Fig. 1a). The octapeptide repeat could also form a helical structure; since the two turn predictions overlap by a single glutamine residue (Fig. 1d), both turns may form. Such a β -turn-rich structure may approximate to a 3_{10} helix, as found in certain alamethicin fragments (Nagaraj & Balaram, 1981).

In native elastin the individual polypeptides are covalently cross-linked via lysine residues, and the elasticity conferred by the β -turns is believed to account partially for the elastic nature of the protein (Urry, 1982). Although C hordein is not elastic, this may be due to the absence of covalent cross-links, the protein containing no cysteine residues (Shewry *et al.*, 1980a, 1981), rather than a lack of intrinsic elasticity in the polypeptide chain.

We are grateful to our colleagues at Rothamsted, notably Dr. B. J. Mifflin, for discussions. We also thank Dr. D. D. Kasarda (U.S. Department of Agriculture, Albany, CA, U.S.A.) and Dr. R. C. Hider (University of Essex, Colchester, Essex, U.K.) for critical comments on this manuscript, and Mr. J. F. March (John Innes Institute, Norwich, U.K.) for the peptide sequencing.

References

- Bush, C. A., Sarkar, S. K. & Kopple, K. D. (1978) *Biochemistry* **17**, 4951–4954
- Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276
- Dufton, M. J. & Hider, R. C. (1977) *J. Mol. Biol.* **115**, 177–193
- Forde, B. G., Kreis, M., Williamson, M. S., Fry, R. P., Pywell, J., Shewry, P. R., Bunce, N. & Mifflin, B. J. (1985) *EMBO. J.* in the press
- Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120
- Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* **8**, 4108–4115
- Isemura, T., Asakura, J., Shibata, S., Isemura, S., Saitoh, E. & Sanada, K. (1983) *Int. J. Peptide Protein Res.* **21**, 281–287

- Ishizaki, H., McKay, R. H., Norton, E. R., Yasunobu, K. T., Lee, J. & Tu, A. T. (1979) *J. Biol. Chem.* **254**, 9651–9656
- Kreis, M., Shewry, P. R., Forde, B. G., Rahman, S. & Miflin, B. J. (1983) *Cell* **34**, 161–167
- Lisowski, M., Siemion, I. Z. & Sobczyk, K. (1983) *Int. J. Peptide Protein Res.* (1983) **21**, 301–306
- Manavalan, P. & Johnson, W. C., Jr. (1983) *Nature (London)* **305**, 831–832
- Miflin, B. J. & Shewry, P. R. (1979) *Seed Protein Improvement in Cereals and Grain Legumes*, vol. 1, pp. 137–158, IAEA, Vienna
- Nagaraj, R. & Balaran, P. (1981) *Biochemistry* **20**, 2828–2835
- Schmitt, J. M. & Svendsen, I. (1980) *Carlsberg. Res. Commun.* **45**, 143–148
- Shewry, P. R. & Miflin, B. J. (1983) in *Seed Proteins: Genetics, Chemistry and Nutritive Value* (Gottschalk, W. & Muller, H., eds.), pp. 143–205, Martinus Nijhoff, The Hague
- Shewry, P. R., Autran, J. C., Nimmo, C. C., Lew, E. J. L. & Kasarda, D. D. (1980a) *Nature (London)* **286**, 520–522
- Shewry, P. R., Field, J. M., Kirkman, M. A., Faulks, A. J. & Miflin, B. J. (1980b) *J. Exp. Bot.* **31**, 393–407
- Shewry, P. R., March, J. F. & Miflin, B. J. (1980c) *Phytochemistry* **19**, 2113–2115
- Shewry, P. R., Lew, E. J. L. & Kasarda, D. D. (1981) *Planta* **153**, 246–253
- Shibata, S., Asakura, J., Isemura, T., Isemura, S., Saitoh, E. & Sanada, K. (1984) *Int. J. Peptide Protein Res.* **23**, 158–165
- Smith, J. A. & Pease, L. G. (1980) *CRC Crit. Rev. Biochem.* **8**, 314–399
- Strickland, E. (1974) *CRC Crit. Rev. Biochem.* **2**, 113–175
- Tanford, C. (1968) *Adv. Prot. Chem.* **23**, 122–282
- Tanford, C. (1970) *Adv. Prot. Chem.* **24**, 1–95
- Urry, D. W. (1982) *Methods Enzymol.* **82**, 673–716
- Urry, D. W., Long, M. M., Ohnishi, T. & Jacobs, M. (1974) *Biochem. Biophys. Res. Commun.* **61**, 1427–1433
- Venkatachalam, C. M. (1968) *Biopolymers* **6**, 1425–1436
- Venkatachalam, C. M. & Urry, D. W. (1981) *Macromolecules* **14**, 1225–1229
- Woody, R. W. (1974) *Peptides Polypeptides Proteins Proc. Int. Symp. 2nd, Rehovot, Israel* 338–350
- Woody, R. W. (1978) *Biopolymers* **17**, 1451–1467