Wheat gluten elasticity: a similar molecular basis to elastin?

Arthur S. Tatham, Peter R. Shewry and Benjamin J. Miflin

Department of Biochemistry, Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ, England

Received 14 September 1984

We have used circular dichroism spectroscopy and structure prediction to study the secondary structure of a group of gluten proteins. They have short α-helices at the N- and C-termini, which are cross-linked by disulphide bonds. The long repetitive central domain has regular β-turns. This structure is similar to that previously proposed for elastin, suggesting a common molecular basis for elasticity.

Wheat *Gluten protein* *Secondary structure* *Elasticity*

1. INTRODUCTION

It is generally accepted that the visco-elasticity of the dough proteins (gluten) is important in determining the baking properties of wheat. However, little is known of the molecular basis of this phenomenon. Gluten is a mixture of at least 50 proteins which are classified into two groups which are either present in aggregates stabilised by covalent disulphide bonds (glutenins) or as monomers associated by non-covalent interactions (gliadins) [1]. Although all of these proteins probably play roles in determining the structure and functionality of gluten, recent studies indicate that the high-M_r subunits of glutenin are particularly important [2,3]. It has been proposed that elasticity is related to the formation of long disulphide-bonded chains of glutenins [4] and the recent demonstrations of cysteine residues close to the N- and C-termini of the high-M_r subunits indicate that they can fulfil this role. There is, however, no information on the forces which may confer elasticity on such polymers. We propose here that the elasticity is due to the presence of an unusual secondary structure which is similar to that previously described for elastin [5]. This suggests that there may be a common secondary structure for elastic proteins of diverse origin.

2. MATERIALS AND METHODS

High-M_r subunits were prepared from wheat cv. Highbury as in [6], but in the presence of 0.1% 2-mercaptoethanol.

To prepare soluble elastin peptides, 100 mg elastin (Sigma) was stirred in 60 ml distilled water. The pH was adjusted to 8.45 with 1 M NH_4OH and 200 mg elastase (Sigma) added. After stirring at 37°C for 8 h the mixture was centrifuged to remove undigested elastin and lyophilised.

CD measurements were made with a Jasco J40CS dichrograph. The result for the high-M_r subunits are calculated using an average monomer M_r of 105.3, calculated from the amino acid composition [6], and 87 for the elastin polypeptides. The units are degree $\cdot$ cm^2 $\cdot$ mol^{-1}. Absorption spectra of the same solutions were recorded with a Cary 210 spectrophotometer. The protein concentration was 1 mg/ml.

The secondary structures were predicted using the methods in [7] and [8]. For the Chou and Fasman analysis predictions were obtained from the products of $N\alpha$ and $N\beta$ of the residue parameters $P\alpha$ and $P\beta$ [9]. Search distances of 6 and 5 were used for helical and sheet regions, respectively. For the β-turn analysis authors in [7] calculated the average probability of a turn occurrence (pt) as 0.55×10^{-4} and selected tetrapeptides

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

with $pt \geq 0.75 \times 10^{-4}$ as probable turns, in this study a cut-off of 1.0×10^{-4} was used. For the method in [8], the unweighted prediction results were used.

3. RESULTS AND DISCUSSION

Although 4–6 high-M_r subunits are present in bread wheat [10] they have similar amino acid compositions, notably 30–35% glutamine, 12–17 mol% proline and 13–21 mol% glycine, and closely related N-terminal sequences [6]. There is also cross-hybridisation between their mRNAs and cDNAs, indicating a high degree of sequence homology [11,12]. The group can therefore be treated on a single protein for physical-chemical characterisation.

Amino acid sequences have been determined directly [6] (fig.1a) and deduced from the nucleotide sequences of cDNA clones (fig.1b,c) [11,12]. Structural prediction, using the methods in [7] and [8], indicated that the first 30 and 38 residues at the N- and C-termini, respectively, are present as α-helices. These contain the only cysteine residues so far detected, two in the N-terminal region (fig.1a) and one in the C-terminal (fig.1c). Although the full extent of the N-terminal helix is not known, the present data show that α-helix contributes only about 8% of the secondary structure of the whole protein, which has been calculated to contain about 650 residues [6].

The rest of the known sequence is composed of interspersed repeated blocks of 6 and 9 residues. A total of 102 β-turns are predicted in this region, although half can be rejected as overlapping with other turns of higher probabilities (table 1). Those with the highest probabilities have Pro.Gly as the two central residues. Thus 204 of the 299 residues in the repeat regions are involved in β-turns which are distributed regularly within the 9 residue blocks and spanning the junctions of the blocks (fig.1).

A second approach to the determination of the secondary structure is CD spectroscopy. Fig.2 shows the far-UV CD spectrum of a preparation of high-M_r subunits dissolved in ethanol–trifluoroethanol, a solvent which promotes ordered hydrogen-bonded conformations. This spectrum shows none of the characteristics associated with the spectra of proteins rich in conventional α-helical or β-sheet conformations [13]; the regular occurrence

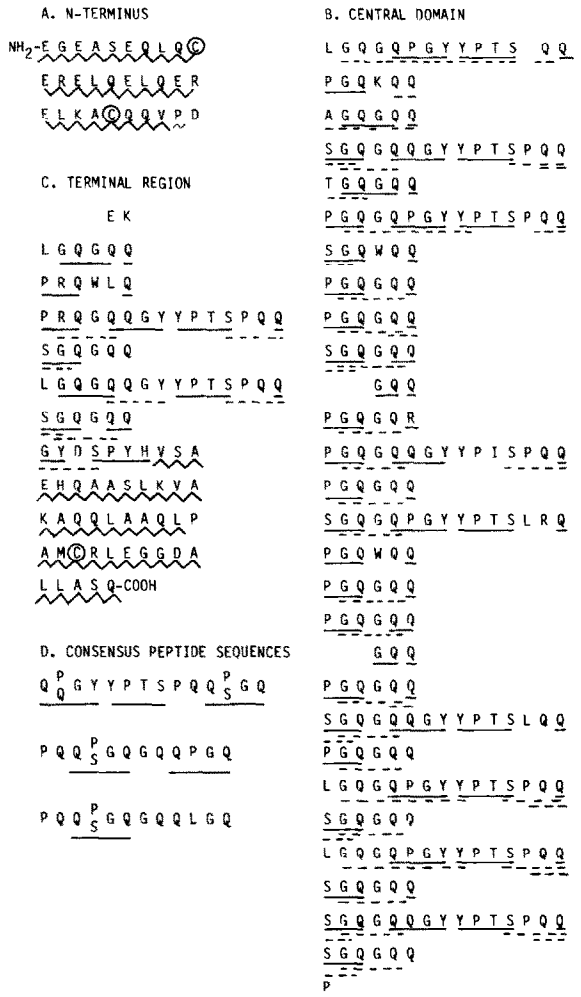

Fig.1. Amino acid sequences of high-M_r subunits and their predicted secondary structures. Where turn predictions overlapped, those with the higher probabilities were accepted and are indicated by solid lines. Those with lower probabilities were rejected and are indicated by broken lines. α-Helical regions are indicated (∿), and regions where the helix may be bent by the presence of a proline residue (~). Cysteine residues are circled. The consensus repeats and their flanking regions are shown in panel D. The sequences were also subjected to the predictive method of authors in [8] which predicted the regions of α-helix at the termini and a mixture of β-turn and random coil for the repeated sequences. Agreement with the two methods for the β-turn predictions was greater than 70%.

of proline residues in the repeated regions would also be expected to preclude the formation of such conformations [7]. It also differs in some respects

Table 1
Predicted β-turns and probabilities from fig.1

Tetra-peptide sequence	Occurrence		Proba-bility ($\times 10^4$)
	High probability (accepted)	Low probability (rejected)	
QPGY	5		5.29
QPGQ	11		4.15
RPGQ	1		3.92
QPRQ	2		2.16
QSGQ	11		1.91
QQGY	4	22	1.86
QQGY	5	1	1.72
QTGQ		1	1.49
YPTS	10		1.70
QQSG		11	1.38
QQGQ	2		1.35
PTSQ		1	1.35
SPQQ		8	1.31
RQGQ		1	1.28
PGYY		5	1.23
QAGQ		1	1.05
Total	51	51	

The turns denoted high probability (accepted) are indicated by solid lines in fig.1. The turns of lower probability were rejected as overlapping with high probability turns. The data in [7,24] showed that proline and serine are the most favoured residues at the second position in the turn, with glycine at the third. It is significant that the majority of the accepted turns contain Pro.Gly and Ser.Gly as the central dipeptide of the turn. Glutamine shows no positional preferences, except at the second where it is unfavourable; the hydrophobic amino acids leucine and isoleucine are not favoured in β-turns, but tyrosine is favoured at the first, third and fourth positions. The residues are predicted in their positional preferences in the high probability β-turns in both the hexapeptide and nonapeptide repeats

from that associated with random coil [14], notably the minimum at 203–204 nm rather than 196–200 nm, and the presence of a shoulder around 230 nm, but consistent with the presence of the predicted β-turns. Theoretical studies and analyses of synthetic peptides have shown that there is no unique CD spectrum associated with β-turns, different turn types giving different spectra [15–18]. Thus the CD spectra of β-turn rich proteins, such as the high-M_r subunits, may represent the sum of a number of spectral types. A similar

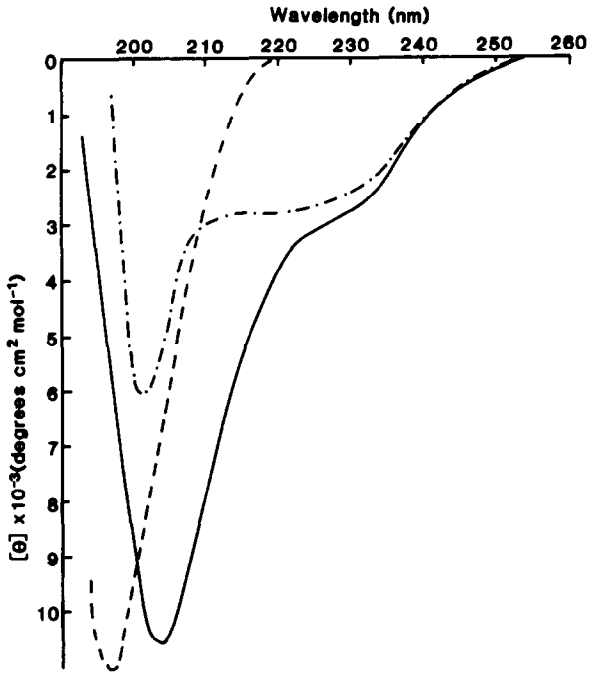

Fig.2. Far-UV circular dichroism (CD) spectra of high-M_r subunits of wheat and elastin peptides. The CD spectrum of the high-M_r subunits (——) was determined in ethanol–trifluoroethanol (7:3, v/v), the subunits were insoluble in aqueous solvents. The elastin peptides were soluble in aqueous buffer (5 mM Tris–HCl, pH 7) (– – –) and in ethanol–trifluoroethanol (7:3, v/v) (-·-·-). The ellipticity (θ) values of both elastin solutions are doubled to facilitate comparison with the high-M_r subunit spectrum.

CD spectrum was found for another β-turn rich protein, the sea anemone toxin anthopleurin A [19].

The structure suggested here for the high-M_r subunits is similar in several respects to that of the elastomeric mammalian connective tissue protein, elastin. This also contains extensive regions of repeated peptides, separated by shorter regions of covalently cross-linked α-helix. The repetitive domains have been shown to be rich in β-turns [20] and also give high predictive probabilities for this conformation (ranging from 1.82 to 5.29×10^{-4}). Furthermore, elastin fragments have similar CD spectra to the high-M_r subunits when dissolved in ethanol–trifluoroethanol [20] (fig.2), although they have a random-coil type spectrum [14] in water (fig.2).

The exact molecular basis for the elastomeric

behaviour of elastin is not known, but it is generally accepted that long polymer chains, a degree of covalent cross-linking and entropic considerations (when a stretched chain represents a state of lower entropy than the unstretched state) are important. One elegant model developed by authors in [21] is based on a β-spiral, a linear helical structure consisting of repeating β-turns. When synthetic β-spirals are covalently cross-linked they form fibrils which behave as an anisotropic elastomer, with an elastic modulus, which dependent on the water content, can be the same as fibrous elastin [20].

Although gluten is not a simple polymer like elastin, the high-M_r subunits may be the major elastic components. We propose that the α-helical N- and C-termini form covalent cross-links via cysteine residues, while the major repeated domain forms elastic fibrils based on β-turns. The fibrils could be stabilized by hydrophobic interactions between the tyrosine residues and by hydrogen bonding between pairs of glutamine residues and between glutamine residues and the peptide backbone. Stretching the protein would disrupt the stable conformation, resulting in an energetically less favourable state. On the removal of the stress the protein would reform the stable conformation.

Aspects of this model probably extend to other elastic proteins. Thus, the giant ($M_r \sim 1 \times 10^6$) elastomeric salivary gland protein of *Chironomus thummi* larvae [22] contains multiple repeats of Pro.Lys.Thr.Ser.Lys.His.Ser.Gly [23]. The method in [7] predicted a high probability of repetitive overlapping β-turns within this sequence ranging from 1.3 to 2.5×10^{-4}, indicating a similar secondary structure to elastin and high-M_r subunits.

ACKNOWLEDGEMENTS

We are grateful to Dr A. Drake (King's College, London) for providing facilities and advice for CD spectroscopy, to Dr M. Sternberg (Birkbeck College, London) for advice and discussion on structural predictions and to Dr J.M. Field for discussions and advice on elastin. The work was supported in part by NATO travel grant 0324/82.

REFERENCES

[1] Miflin, B.J., Field, J.M. and Shewry, P.R. (1976) in: Seed Proteins (Daussant, J. et al. eds) pp.255–319, Academic Press, London.
[2] Field, J.M., Shewry, P.R. and Miflin, B.J. (1983) J. Sci. Food Agric. 34, 370–377.
[3] Payne, P.I., Corfield, K.G., Holt, L.M. and Blackman, J.A. (1981) J. Sci. Food Agric. 32, 51–60.
[4] Ewart, J.A.D. (1977) J. Sci. Food Agric. 28, 191–199.
[5] Venkatachalam, C.M. and Urry, D.W. (1981) Macromolecules 14, 1225–1229.
[6] Shewry, P.R., Field, J.M., Faulks, A.J., Parmar, S., Miflin, B.J., Dietler, M.D., Lew, E.J.-L. and Kasarda, D.D. (1984) Biochim. Biophys. Acta 788, 23–34.
[7] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251–276.
[8] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[9] Dufton, M.J. and Hider, R.C. (1977) J. Mol. Biol. 115, 177–193.
[10] Holt, L.M., Astin, R. and Payne, P.I. (1981) Theor. Appl. Genet. 602, 237–243.
[11] Forde, J., Forde, B.G., Kreis, M., Fry, R., Shewry, P.R. and Miflin, B.J. (1983) FEBS Lett. 162, 360–366.
[12] Thompson, R.D., Bartels, D., Harberd, N.P. and Flavell, R.B. (1983) Theor. Appl. Genet. 67, 87–96.
[13] Manavalan, P. and Johnson, W.C. jr (1983) Nature 305, 831–832.
[14] Dearborn, D.G. and Wetlaufer, D.B. (1970) Biochem. Biophys. Res. Commun. 39, 314–320.
[15] Woody, R.W. (1974) in: Peptides, Polypeptides and Proteins, Proceedings 2nd International Symposium (Blout, E.R. et al. eds) pp.338–350, Wiley, New York.
[16] Bush, C.A., Sarkar, S.K. and Kopple, K.D. (1978) Biochemistry 17, 4951–4954.
[17] Lisowski, M., Siemion, I.Z. and Sobczyk, K. (1983) Int. J. Peptide Protein Res. 21, 301–306.
[18] Urry, D.W., Long, M.M., Ohnishi, T. and Jacobs, M. (1974) Biochem. Biophys. Res. Commun. 61, 1427–1433.
[19] Ishizaki, H., McKay, R.M., Norton, T.R., Yashunobu, K.T., Lee, T. and Tu, A.T. (1979) J. Biol. Chem. 254, 9651–9656.
[20] Urry, D.W. (1982) Methods Enzymol. 82, 673–716.
[21] Urry, D.W., Venkatachalam, C.M., Long, M.M. and Prasad, K.U. (1983) in Conformation in Biology (Srinivason, R. and Sarma, R.M. eds) pp.11–27, Adenine Press, New York.
[22] Serfling, E., Meyer, L., Rudolph, A. and Steiner, K. (1983) Chromosoma 88, 16–23.
[23] Baumlein, H., Wobus, U., Gerbi, S. and Kafatos, F.C. (1982) EMBO J. 1, 641–647 (1982).
[24] Chou, P.T. and Fasman, G.D. (1977) J. Mol. Biol. 115, 135–175.