Structural characterization of the central repetitive domain of high molecular weight wheat gluten proteins.

Swieten, Eric van

The interest of studying HMW proteins

Wheat gluten proteins enable dough to be expanded, by yeast fermentation. This special property comes from the microscopically fine network consisting of proteins, that is being formed during the kneading of dough. It results in a substance that is viscous and elastic at the same time. In physical terms this combination of two properties is called viscoelasticity. The proteins that make the strongest contribution to the viscoelasticity are also the largest proteins occurring in wheat. They are called high molecular weight proteins, or HMW proteins in short. Although many studies have been performed, a thorough understanding of the spatial structure of these proteins is lacking. A better insight in the structure of HMW proteins could add to a better understanding of the viscoelastic properties of dough.

Original objectives and starting-points for research

The general objective of these studies was to get a description of the spatial structure of the HMW protein Dx5. In first instance, structural determination using nuclear magnetic resonance (NMR) spectroscopy was emphasized. Dx5 was chosen because of the availability of the gene, so that certain parts of the protein could be studied separately. Apart from that, Dx5 is known for its highly favorable elastic properties in comparison with the more than 20 other HMW proteins currently known. The molecular genetic approach also offered the possibility to produce Dx5 in other organisms, in order to obtain pure protein in a relatively simple fashion. We chose to express Dx5 and parts of it in E. coli, using a histidine-tag coupled to the C-termini. Using this His-tag, one-step purification could be executed by means of affinity chromatography.

HMW proteins consist of three domains: a central repetitive domain flanked by two smaller, terminal, domains. The N-terminal domain and a part of the central domain were expressed in E. coli, and purified for structural characterization. The C-terminal domain was obtained using preparative protein synthesis. Besides a high purity, the most important starting-point for structural studies was to get the three domains in a soluble state. The soluble state was needed, given the original objective to use NMR for high-resolution structural determination. Studying the HMW proteins in solution also offered the possibility to create a homogeneous environment, in contrast to partially aggregated systems in liquid, or systems in solid state at surfaces. Information on HMW's under heterogeneous circumstances was not regarded as useful, because it has not been established to what extent such conditions represent the actual situation in dough or bread.

Interim evaluation and adapted objectives for further research

The original objectives were adjusted according to the results obtained. The N-terminal domain appeared to be non-soluble under the desired experimental conditions for NMR, making further studies using NMR not viable. Initial

characterization of the secondary structure was carried out in 0.1 % SDS, in which it was only soluble at low concentrations. The C-terminal domain was soluble in water, but did not have a well-defined structure. Determination of the uniquely folded structure of the C-terminal domain using NMR was therefore not relevant. The most important findings in the terminal domain studies are included in chapter 1.
NMR was, *a priori,* not useful for characterization of the repetitive domain, due to the size of this domain and the overlap of NMR resonances. It was nevertheless successfully applied for the determination of the beta-turn structure in peptides, derived from the repetitive sequence (chapter 1). These were the starting points for further examination of the beta-turn structure using other, lower-resolution, techniques (chapters 2 and 3). Former studies indicated that the repetitive domain did not possess a compact three-dimensional structure. Therefore the size and shape of the repetitive domain were examined (chapters 4 and 5). For these purposes spectroscopic and scattering experiments were carried out on the cloned dB1 protein, that roughly takes up a quarter of the full repetitive domain in HMW Dx5. Similar studies were carried out on dB4, an oligomer consisting of four sequentially cloned dB1 copies. The amount of amino acids present in dB4 is approximately equal to the amount in the full repetitive Dx5 domain.

Solubility and aggregation of HMW domains
The repetitive domain is highly soluble in aqueous solutions, in contrast to the moderate to low solubilities of the terminal domains. This implies that the solubility of the full repetitive domain is governed by the solubilities of the terminal domains. The indication for a water-layer around dB1 and dB4 according to conclusions from the hydrodynamic work was in agreement with the high solubilities and the hydrogen-bonding ability of the repetitive sequence. The assurance that size and shape of dB1 and dB4 molecules were obtained, and not corresponding values of partially aggregated systems, was provided by testing the aggregation state in several buffers using dynamic light scattering (DLS). Aggregation was controlled by use of carboxylate buffers. These buffers probably form hydrogen bonds with the glutamines in HMW proteins, thus competing with mutual hydrogen bonds between HMW's which would lead to aggregation. For the first time, scattering and hydrodynamic experiments were performed on the repetitive domain under monodisperse conditions.

The beta-turn structure of the repetitive HMW domain
Circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopy showed that dB1 is rich in beta-turns under aqueous conditions. The results have indicated random coil structure in addition to the beta-turn structure, which is in good agreement with previous studies on HMW proteins. The amino acid composition of the central domain strongly deviates from the average composition in proteins, with

41% glutamine, 23% glycine and 14% proline. The basis for this peculiar composition lies in the repetitive character of the central domain that consists of alternating stretches of mainly 6 or 9 amino acids (hexa- and nona-repeats). It is known from other studies that peptides with this composition adopt beta-turn structures. Beta-turns are short structural elements with a length of 4 amino acids, which effect reversal of the direction of the protein chain. They are common structural building blocks in proteins, in addition to alpha-helical and beta-sheet structures.

The stability of beta-turn structure of the repetitive HMW domain

What happens with the beta-turn structure in experimental environments different from water at room temperature? To examine this, a cloned part of the central domain of HMW Dx5 (dB1) was produced in E. coli and purified. Structural changes in dB1 in the soluble state were followed using CD and FTIR spectroscopy. Subsequently the influence of temperature, the detergent sodium-dodecyl-sulfate (SDS) and the organic solvent trifluoroethanol (TFE) were tested on the secondary structure. The studies implied a high stability of the beta-turn structure in the repetitive HMW domain as function of both temperature and TFE. A distinct increase in the aggregation of dB1 was observed in the range 0.5-2.0 % SDS, with a concomittant decrease of β-turn structure. At higher SDS concentrations, this process seemed to reverse.

Statistical analysis of beta-turn probability in different HMW proteins

Structural characterization with NMR in particular has led to the identification of beta-turns at specific locations in the hexa- and nona-repeats in previous work. That is the reason why beta-turns can be assigned in the full repetitive domain to the corresponding hexa- and nona-repeats. In this way, 80-85 % of the repetitive sequence is covered with beta-turns. The consensus sequences of the beta-turns can be deduced directly from the consensus sequences of the hexa- and nona-repeats. The high statistical probability that these sequences in HMW proteins form beta-turns, implies that beta-turn structure is favored above other types of structure. This was further supported by the lower variation at the terminal positions of each beta-turn in comparison with the two central positions. It suggests that conservation of the amino acids in the terminal position is important, which agrees with the participation of the amino acids at these positions in structure-stabilizing hydrogen-bonds of the beta-turns.

Size and shape of the repetitive domain

The cloned dB1 and dB4 proteins were examined in soluble state using small-angle neutron scattering (SANS). In addition, two complementary hydrodynamic techniques, analytical ultracentrifugation (AU) and dynamic light scattering (DLS), were used for characterization. The global spatial structure determined with SANS of

both dB1 and dB4, was in agreement with a worm-like chain (WLC), a model that is frequently used in polymer theory. The model describes a flexible cylinder with a length of 235 and 900 Å, for dB1 and dB4 respectively, and a cross-section of approximately 15 Å for both proteins. The results imply a longer repetitive domain compared to previous studies on HMW's in solution. The hydrodynamic results confirm the asymmetrical shape of dB1 and dB4, and were in good agreement with the WLC model.

Flexibility of the repetitive domain and elasticity

The flexibility of both proteins is characterized by a persistence length of about 13 Å. Their structures are identical, which implies that the central HMW domain can be elongated while retaining its structural characteristics. However, compared to the previously proposed structure of a stiff rod, our experiments have clearly indicated more flexibility of the cylinder. The flexibility introduces a new mechanism for elasticity, and enables the central domain to bend to such an extent that the terminal HMW domains are able to form intramolecular disulfide bonds. This phenomenon has already been demonstrated in previous studies on HMW proteins.

Beta-spiral as model for the repetitive domain

It seems conceivable that the flexible cylinder results from a helical structure, which resembles the beta-spiral suggested in elastin. Elastin is present in large quantities in human tissue and has been proposed to confer elasticity and flexibility on these tissues. It possesses a large number of proline and glycine residues which are responsible for a very high content of beta turns. This has led investigators to suggest a spiral-like structure in elastin, which is called the beta-spiral. HMW proteins have been compared with elastin because of the repetitive character and the high proline and glycine contents in the central domain. Elastin has elastic properties, as the name indicates. Similar characteristics could explain the functional properties of HMW proteins in dough.

The structure of the repetitive domain in relation to bread-making quality

To what extent does knowledge on the molecular structure of HMW proteins generated in this thesis add to the understanding of bread-making quality? Physical aspects such as water-binding properties and viscoelasticity determine this quality. While it should be clear that an extrapolation of the features of a pure system in solution, to a heterogeneous dough, is not straightforward, a number of our observations on HMW proteins can be linked to the functional properties of dough and bread:

- The observed stability of the beta-turn structure is in line with the theory that the gluten network is well-resistant to heating during the baking process.

- The solubility of the central domain is much higher than the solubilities of the terminal domains. This difference emphasizes the dual role of HMW proteins in water-binding and aggregation. The high solubility and water-layers around dB1 and dB4 showed that the repetitive domain strongly interacts with water. The experimental link between the good solubility and intermolecular hydrogen bonding was strongly suggested by the monomeric state of dB1 in carboxylate-based buffers in contrast to the formation of aggregates in phosphate buffer.
- Dough strength and extensibility are mediated by disulfide bonds. The interplay of covalent and non-covalent interactions defining the branching points, forms one element for a consistent and viscoelastic dough. Another element is the length or size of the chains bridging the branching points. The flexibility of the repetitive structures as defined in this study, is possibly also present in dough networks, and might add to the viscoelasticity of dough.
- The statistical work does not give a clear indication for large differences in properties between the repetitive domains of different HMW proteins.

This thesis has provided a better insight in the molecular structure of the repetitive HMW domain. Advances were also made towards a further understanding of the functional behaviour of wheat gluten HMW proteins. Stepwise addition of more components to the system is now needed to come to a full description of the complex role of HMW proteins in dough and bread.