

## IMPROVING THE PROTEIN CONTENT AND QUALITY OF TEMPERATE CEREALS: WHEAT, BARLEY AND RYE

**Shewry, P. R.**

*Long Ashton Research Station, University of Bristol, UK*

**Keywords:** Rye, barley, wheat, protein content, protein quality.

### Contents

1. Introduction
  2. Increasing Total Grain Protein
  3. Improving Grain Amino Acid Composition
  4. Conclusion
- Acknowledgements  
Glossary  
Bibliography  
Biographical Sketch  
To cite this chapter

### Summary

Wheat, barley and rye are major sources of protein for the nutrition of humans and livestock, but are deficient in certain essential amino acids when used as food for monogastric animals. In particular, they contain low levels of lysine (the first limiting amino acid) and to a lesser extent, threonine (the second limiting amino acid) resulting from deficiencies of these amino acids in the prolamin storage proteins, which account for about half of the total nitrogen in the mature grains. The total protein content of the grain can be manipulated by adding fertilizer nitrogen but “genes” conferring high grain protein have been identified in wild tetraploid wheats. However, these have not yet been exploited to increase the protein content of cultivated wheat. Similarly, mutant genes have been identified, which confer the high lysine phenotype to barley grain, resulting from decreases in the proportion of lysine-poor prolamins and/or increases in specific lysine-rich proteins. However, these genes are also associated with low yields, and have not been successful when incorporated into cultivated lines. Genetic engineering has great potential for increasing the essential amino acid content of temperate cereals, using one of two strategies. Transformation with bacterial genes encoding feedback insensitive enzymes of lysine biosynthesis may lead to increased accumulation of lysine as free amino acid. Alternatively, expression of additional genes for lysine-rich proteins may lead to increased accumulation of lysine in proteins. These two strategies may also be combined.

### 1. Introduction

Barley, wheat and rye are related members of the Tribe Triticeae and consequently share many genetic and biochemical characteristics. Cultivated wheat exists in diploid, tetraploid and hexaploid forms but the vast majority of the wheat grown worldwide is of two types – hexaploid *Triticum aestivum* var *aestivum* (genome constitution AABBDD)

(bread wheat) and tetraploid *T. turgidum* var *durum* (genome constitution AABB) (durum or pasta wheat). However, other varieties of these two species have been cultivated historically and are still grown in small amounts in some parts of the world, as are other primitive diploid and tetraploid species. In addition, a number of wild diploid species related to the progenitors of the A, B and D genomes of cultivated wheats and wild *T. turgidum* var. *dicoccoides* also occur, providing a reservoir of valuable variation for plant breeding.

In contrast to wheat, cultivated barley comprises only one species, *Hordeum vulgare*, which is diploid and crosses readily with the wild form *H. spontaneum*, which is now generally considered to belong to the same species. Whereas barley and wheat are inbreeding, rye (*Secale cereale*) is an out-breeding diploid. A number of wild species occur, but cross-fertility between these may be limited by the presence of chromosome translocations.

Wheat, barley and rye are important sources of protein for the nutrition of humankind and of livestock, which is in turn used to provide dietary protein. The content and nutritional quality of the grain proteins is therefore an important consideration (see *Plant Based Sources of Proteins and Amino Acids in Relation to Human Health*). However, it is also important to bear in mind the impact of the grain proteins on the processing properties, with most cereals being consumed by humans after processing into food or beverages rather than as whole or milled grain.

When used for animal feed, the quality requirements depend on the animal in question (particularly whether ruminant or non-ruminant) and the stage of development. In the case of ruminants, the microflora present in the rumen can synthesize all twenty protein amino acids, which is sufficient to provide for the dietary requirements of the animal, except under high production conditions. For non-ruminants, about half of these amino acids are essential, in that they must be provided in the diet. Thus, if even one essential amino acid is present in insufficient amounts, the remaining amino acids will be broken down, leading to nitrogen loss (and subsequent environmental pollution in high production conditions) and poor growth. When compared with the WHO requirements of essential amino acids for humans, wheat, barley and rye are seen to be deficient in lysine, with threonine being the second limiting amino acid (Table 1).

Amino acids	Wheat	Rye	Barley	WHO Recommended levels
Cysteine	2.6	2.9	2.9	} 3.5
Methionine	1.3	1.7	1.7	
Lysine	2.0	3.3	3.1	5.5
Isoleucine	3.6	3.6	3.6	4.0
Leucine	6.7	6.7	7.2	7.0
Phenylalanine	5.1	4.9	5.5	} 6.0

Tyrosine	2.6	2.1	2.7	4.0
Threonine	2.7	3.4	3.3	1.0
Tryptophan	1.1	1.8	2.0	5.0
Valine	3.7	4.4	4.6	-
Histidine	2.2	2.1	1.9	-

Amounts are expressed as g/100g protein. Cysteine and tyrosine are not truly essential as they can be synthesised from methionine and phenylalanine, respectively. Hence combined values for cysteine + methionine and phenylalanine + tyrosine are given by WHO. Similarly, histidine is essential for human children but not adults. Data taken from J.A.D. Ewart (1967) Amino acid analyses of cereal flour proteins. *Journal of the Science of Food and Agriculture* **18**, 548-552 and FAO (1973) Energy and Protein Requirements, FAO Nutritional Meeting Report, Series No.52, WHO Technical Report Series No.552, Rome.

Table 1. Contents of essential amino acids in grain of wheat, rye and barley compared with the WHO recommended levels.

Much of the wheat consumed by humans is in the form of bread, which occurs in a vast range of forms in different cultures, including leavened pan and hearth-baked breads, flat and pocket breads and steamed breads, with numerous variants within these different types. In addition, a range of types of noodles (made from bread wheat) and pasta (from durum wheat) and other baked products (cakes, cookies etc.) are made. Consequently, it is important to maintain the quality for specific end uses when attempts are made to improve wheat protein quality – failure to do this may lead to lack of acceptance irrespective of any improvement in nutritional quality.

Barley and rye are consumed less frequently in processed foods than wheat, with rye flour often being blended with wheat flour to improve its functional properties. The major “food” use of barley in developed countries is for malting, brewing and distilling. In this case, processing quality is the overwhelming consideration. However, barley is still a staple food in some other parts of the world and nutritional quality is, therefore, important.

### 1.1. Cereal Grain Proteins

The systematic study of plant proteins, including those of cereal grains, is founded on the work of T.B. Osborne, working at the Connecticut Agricultural Research Station, from 1886 until 1928. He classified cereal proteins into four groups, based on their sequential extraction in a series of solvents. Thus, the albumins were extracted in water followed by the globulins in sodium chloride solution, the prolamins in aqueous alcohols and the glutelins in dilute acid or alkali. This system has proved remarkably durable and still provides a framework for modern cereal chemistry. However, it is inevitable that the extraction procedures have been modified in the light of our improved knowledge of the structures, functions and genetics of the proteins present within the fractions.

We now know that the albumin and globulin fractions of cereals contain predominantly structural, metabolic and protective proteins, although 7S storage globulins are present in the aleurone layer of the endosperm and the scutellum of the embryo. In contrast, the prolamins comprise the major grain storage proteins and are located in the starchy endosperm cells. However, whereas prolamins were classically extracted with aqueous (60-70 per cent (v/v)) ethanol, it is now usual to use other alcohols which give more efficient extraction (often 50 per cent (v/v) propan-1-ol) and to include a reducing agent (2-mercaptoethanol or dithiothreitol) to extract prolamins which are not soluble in aqueous alcohols due to their assembly into high  $M_r$  polymers stabilized by inter-chain disulphide bonds. Once these proteins are removed, the residual glutelin proteins, which probably include insoluble cell wall proteins and other structural proteins, are usually extracted using a solvent such as sodium dodecylsulphate (SDS) rather than acid or alkali, which may result in some degradation. Finally, even after the extraction of glutelins, a small proportion of the total proteins remains in the residual meal.

When extracted efficiently, the prolamins of wheat, rye and barley may account for 50 per cent or more of the total grain nitrogen, as discussed below. However, these fractions contain only about 1 mol percent or less lysine compared with about 4-5 mol percent lysine in the albumin, globulin, glutelin and residual fractions (Table 2). Similarly, the prolamins are also low in threonine. Consequently, it can be concluded that the poor nutritional quality of the whole grain proteins results from deficiency of essential amino acids in the prolamins storage proteins. It is, therefore, necessary to briefly describe the characteristics of the prolamins before discussing strategies to manipulate their amount and composition.

Amino acid	Barley	Wheat	Rye
Lysine	1.0	0.9	0.9
Histidine	1.2	1.7	1.3
Arginine	2.6	2.1	1.5
Aspartate <sup>a</sup>	1.9	3.0	2.1
Threonine	2.5	2.7	2.5
Serine	2.9	5.4	5.8
Glutamate <sup>a</sup>	31.8	32.6	35.8
Proline	20.1	17.3	20.2
Cysteine	2.9	1.9	2.5
Glycine	3.2	5.6	4.2
Alanine	3.0	3.7	2.7
Valine	5.1	3.8	4.4
Methionine	0.6	1.2	1.1
Isoleucine	4.1	3.7	3.0
Leucine	7.4	7.1	5.7
Tyrosine	2.6	2.6	1.7
Phenylalanine	5.2	4.5	4.6

<sup>a</sup>Values reported for aspartate and glutamate include asparagine and glutamine, respectively.

Tryptophan was not determined.

Data taken from Kreis, M., Shewry, P.E., Forde, B.G., Rahman, S., Bahramian, M.B. and Mifflin, B.J. (1984) Molecular analysis of the effects of the mutant *lys 3a* gene on the expression of *Hor* loci in developing endosperms of barley (*Hordeum vulgare*). *Biochemical Genetics* **22**, 231. Byers M., Mifflin, B.J. and Smith, S.J. (1983) A quantitative comparison of the extraction of protein fractions from wheat grain by different solvents, and of the polypeptide and amino acid composition of the alcohol-soluble proteins. *Journal of the Science of Food and Agriculture* **34**, 447-462 and Bright, S.W.J and Shewry, P.R. (1983) Improvement of protein quality in cereals. *CRC Critical Reviews in Plant Sciences* **1**, 49-93.

Table 2. Amino acid compositions (mol per cent) of total prolamins fractions from barley, wheat and rye.

## 1.2. Cereal Prolamins

The prolamins fractions of wheat, rye and barley are complex mixtures of proteins, which vary in their composition between different genotypes of the same species. The total number of components has not been determined precisely, but analyses by one- and two-dimensional electrophoresis show the presence of at least 50 individual proteins in hexaploid bread wheat and 20-30 in barley and inbred lines of rye (Figure 1). The individual components are classically divided into groups based on their solubility and electrophoretic properties, the groups being given different names in the three species (hordeins, secalins and gliadins + glutenins in barley, rye and wheat, respectively) (Figure 1). However, the determination over the past two decades of the complete amino acid sequences of many individual prolamins has allowed a more precise classification to be proposed, based on amino acid sequence relationships. This recognizes three major groups, called the sulphur-rich (S-rich), S-poor and high molecular weight (HMW) prolamins, which differ in their  $M_r$  and amino acid compositions, including contents of the essential amino acid lysine, threonine, cysteine and methionine (Table 3).

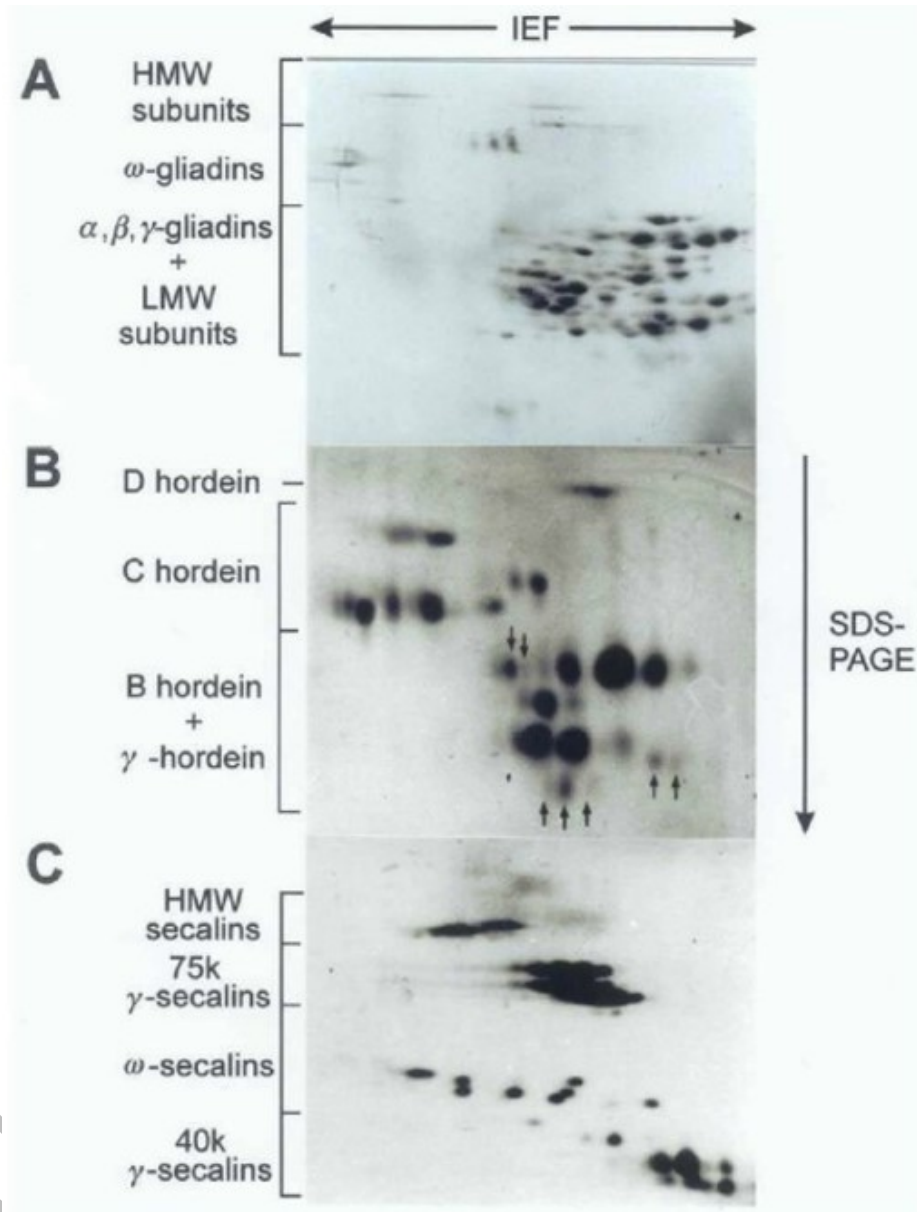


Figure 1. Two-dimensional analysis (IEF/SDS-PAGE) of reduced and pyridylethylated total prolamins from A, wheat (cv. Chinese Spring); B, barley (cv. Carlsberg II) and C, rye (cv. inbred line MPI 109). The arrows in part B indicate  $\gamma$ -hordeins, the other S-rich prolamins being B hordeins.

PROLAMIN TYPE	% TOTAL FRACTION	COMPONENT GROUPS			POLYMER FORMATION	$M_r$	PARTIAL AMINO ACID COMPOSITION (mol%)
		BARLEY	WHEAT	RYE			
HMW Prolamins	2-10	D hordein	HMW subunits	HMW secalins	Yes	65-90 000	≅1% Lys 3-7% Thr 0.5-1.5% Cys 0.2-0.5% Met
S-poor Prolamins	10-20	C hordeins	ω-gliadins	ω- secalins	No	30-80 000	0-0.5% Lys 1-2% Thr 0% Cys 0-0.2% Met
S-rich Prolamins	70-80	γ-hordeins B hordeins	γ-gliadins α-gliadins LMW subunits	γ-secalins	No No Yes	30-55 000	0.5-1% Lys 2-3% Thr 2-3% Cys 1-2% Met

Table 3. Summary of the types and properties of prolamins present in wheat, barley and rye.

The S-rich prolamins are the quantitatively major group in all three species and also contain higher levels of lysine, cysteine and methionine than the S-poor and HMW prolamins. However, higher levels of threonine are present in the HMW prolamins, particularly in D hordein, which contains about 8 mol percent threonine. In contrast, the C hordeins contain the lowest amounts of essential amino acids with some components completely lacking lysine, cysteine and methionine. It should be noted, however, that with the exception of threonine in some HMW prolamins and of cysteine + methionine in S-rich prolamins, the proportions of lysine, threonine, cysteine and methionine present in the various prolamins groups are substantially lower than the levels recommended by the WHO. Consequently, there would be limited impact on the nutritional quality of the whole grain if the proportions of the prolamins groups were changed.

The groups of prolamins also differ in their ability to form disulphide-stabilized polymers. This is particularly important in relation to the functional properties (i.e., processing) of wheat. In wheat, the prolamins are the major components of the gluten proteins, which form a continuous network in dough. This network is responsible for conferring the cohesiveness and visco-elastic properties, which allow wheat dough to be processed into bread, noodles and pasta and a range of other foods. Gluten proteins are classically divided into two fractions, the gliadins which are monomeric and the polymeric glutenins. The gliadins comprise the S-poor  $\omega$ -gliadins and the S-rich  $\alpha$ -type and  $\omega$ -type gliadins while the glutenins can be separated, after reduction of inter-chain disulphide bonds, into the HMW subunits and the S-rich LMW subunits of glutenin. Although this gliadin/glutenin classification is not biologically valid, in that both fractions contain S-rich prolamins, it remains widely used because the fractions have functional importance, with the gliadins being largely responsible for gluten viscosity and the glutenins for elasticity. As the functional properties of wheat gluten underpin almost all aspects of grain utilization, it is important that they should not be affected by any attempts to improve the nutritional quality for human consumption by genetic manipulation.

## **2. Increasing Total Grain Protein**

The protein content of wheat, barley and rye is, of course, important in relation to the nutritional quality for humans and for livestock feed. However, it also has implications for the processing quality of wheat and barley. High protein contents are required for wheat used for baking pan breads and blending, above 13 percent for Canadian Western Red Spring Wheats and 11-13 percent for UK wheats, with lower levels required for other types of bread and for noodles and other food uses. In the case of barley, low protein levels (generally below about 11 per cent in the UK and Europe) are required for malting, brewing and distilling, with higher levels resulting in poorer quality.

The protein content of the grain appears to be determined by a combination of genetic and environmental factors, with variation in nutrition (particularly availability of N) resulting in considerable variation in protein content within genetically determined limits. Due to this strong environmental influence, it is difficult to compare values reported for the protein content of grain grown in different locations or in different years. Thus, although it is generally considered that wheat has a lower protein content



than barley and rye, it is difficult to substantiate this with precise figures. For example, a range of 15-17 percent protein has been reported for 23 samples of hard red winter wheats grown in the USA, while protein contents of 6.5 to 14.5 percent have been reported for rye with the higher values for grain grown in N. America. A similarly wide range of 8.5 to 21.2 percent with a mean of 13.1 percent was reported for fourteen hundred lines of barley. The protein content of barley has also been reported to vary from 8.1 to 14.7 percent when the same genotype was grown in different locations while a narrower range (about 1.3 – 2.0 percent N corresponding to about 7.2 – 11.5 percent protein) was reported in a single variety grown in a replicate field trial in which N fertilization and cropping history were varied (Figure 2).

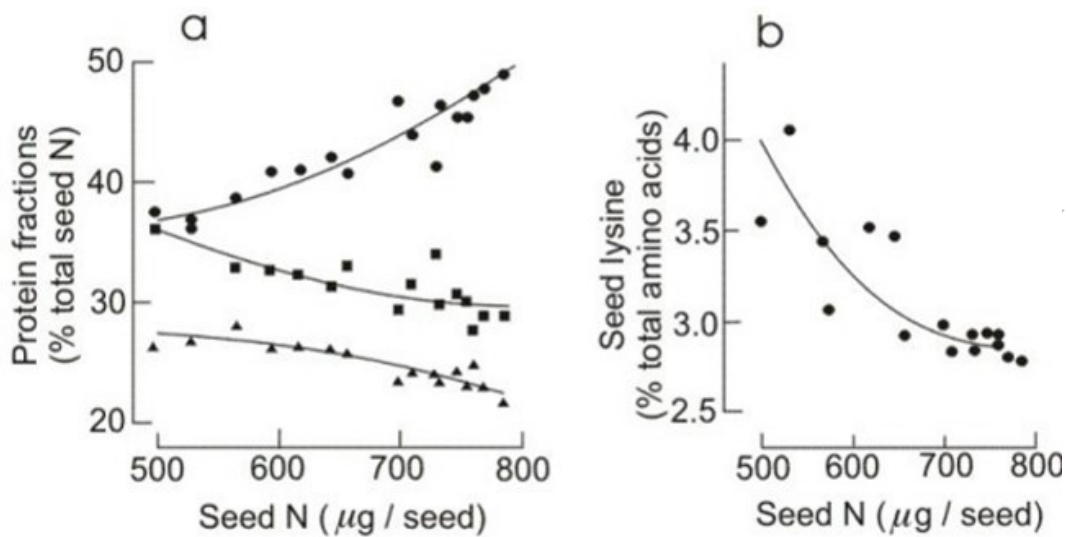


Figure 2. The relationship between total grain N and the relative amounts of protein fractions (part a) and lysine (part b) in barley. Taken from Kirkman M.A., Shewry, P.R. and Mifflin, B.J. (1982) The effect of nitrogen nutrition on the lysine content and protein composition of barley seeds. *Journal of the Science of Food and Agriculture* **33**, 115-127, with permission of the publishers.

Variation in the total protein content of the grain results in specific effects on the protein composition, which in turn affect the contents of lysine and other essential and non-essential amino acids. The main impact of increased availability of nitrogen is an increase in the prolamin storage proteins, which act as a sink for any nitrogen which is in excess of that required for the synthesis of structural and metabolic components. This is illustrated for field-grown barley in Figure 2, showing that hordein increases from 36 to 49 percent of the total nitrogen as the latter increases from about 1.3 to 2.0 percent dry weight. Even more extreme effects can be observed in material grown in the glasshouse with about 60 percent hordein present in grain with almost 4 percent nitrogen. Furthermore, increases in grain N may also result in effects on the composition of the prolamin fractions, with an increase in the proportion of the sulphur-poor C hordeins occurring in barley (from about 13 to 18.5 percent of the total hordein fraction in the material shown in Figure 2). Similar increases in the proportions of S-poor C hordeins and  $\omega$ -gliadins occur in barley and wheat, respectively, grown under conditions of sulphur deficiency, indicating that high nitrogen availability may result in

limitation in sulphur-containing amino acids inside the developing grain, unless sulphur is also provided.

Prolamins are low in lysine, with significantly lower levels in the S-poor than in the S-rich and HMW prolamins. Consequently, significant decreases occur in the proportions of lysine present in the total grain proteins, and hence the nutritional quality is affected, when barley (Figure 2) or wheat is grown with high nitrogen availability. As these effects are broadly reproducible with different genotypes and in different environments, it is possible to make approximate calculations of the lysine contents of grains with different nitrogen levels using regressions, which are species-dependant (Figure 3).

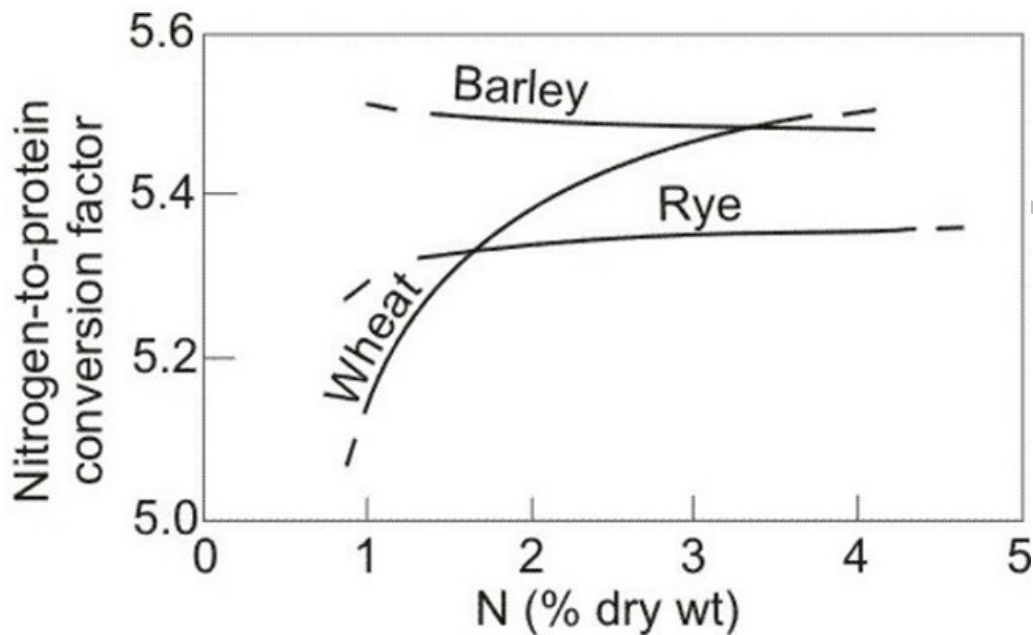


Figure 3. The relationship between total grain N and the N to protein conversion factor in seeds of wheat, barley and rye. Redrawn from Mossé, J. (1990) Nitrogen to protein conversion factor for ten cereals and six legumes or oilseeds. A reappraisal of its definition and determination. Variation according to species and to seed protein content. *Journal of Agriculture and Food Chemistry* **38**, 18-24, with permission of the publishers.

As prolamins contain high levels of glutamine, a nitrogen-rich amino acid, the nitrogen to protein conversion ratio also varies with total grain N in a species-dependent fashion (Figure 4).

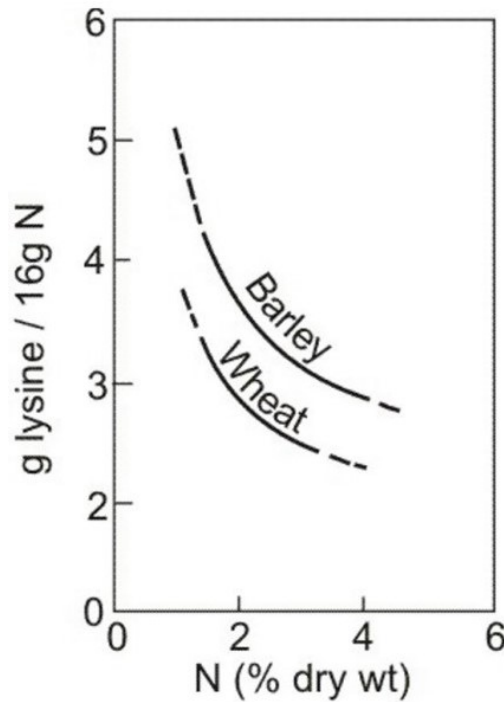


Figure 4. The relationship between total grain N and the % lysine in seeds of barley and wheat. Redrawn from Mossé, J. and Huet, J.C. (1990) Amino acid composition and nutritional score for 10 cereals and six legumes or oilseeds: causes and ranges of variations according to species and to seed nitrogen content. *Sciences des Aliments* **10**, 151-173 with permission of the publishers.

There is, in general, a negative correlation between grain protein content and yield. Since yield is the major determinant of profit for the farmer, there has been limited interest in developing high protein cultivars. The relationship between protein and yield based on data from 74 experiments on wheat (including 13 from Europe), 18 experiments on barley and 14 on other cereals (there being no data for rye) is shown in Figure 5. It can be concluded that a negative relationship exists between yield and protein content, but there is no simple physiological explanation for this.

Despite this inverse relationship between yield and protein content, the importance of grain protein for the processing of wheat has led to a number of plant breeding programmes focused on high protein grain, particularly in the USA and Canada. These programmes have exploited either naturally occurring high protein lines of wheat or lines of wild tetraploid *T. turgidum* var *dicoccoides*. High protein cultivars used as parents for plant breeding include Atlas 50 and 66, Nap Ha 1, Minnpro and Plainsman V for bread wheat and Trinakria for durum wheat. In Plainsman V, the high protein gene(s) may be derived from a cross with *Aegilops ovata* (oat grass) and cultivars based on this material with increases in grain protein and competitive yields have been released in N. America.

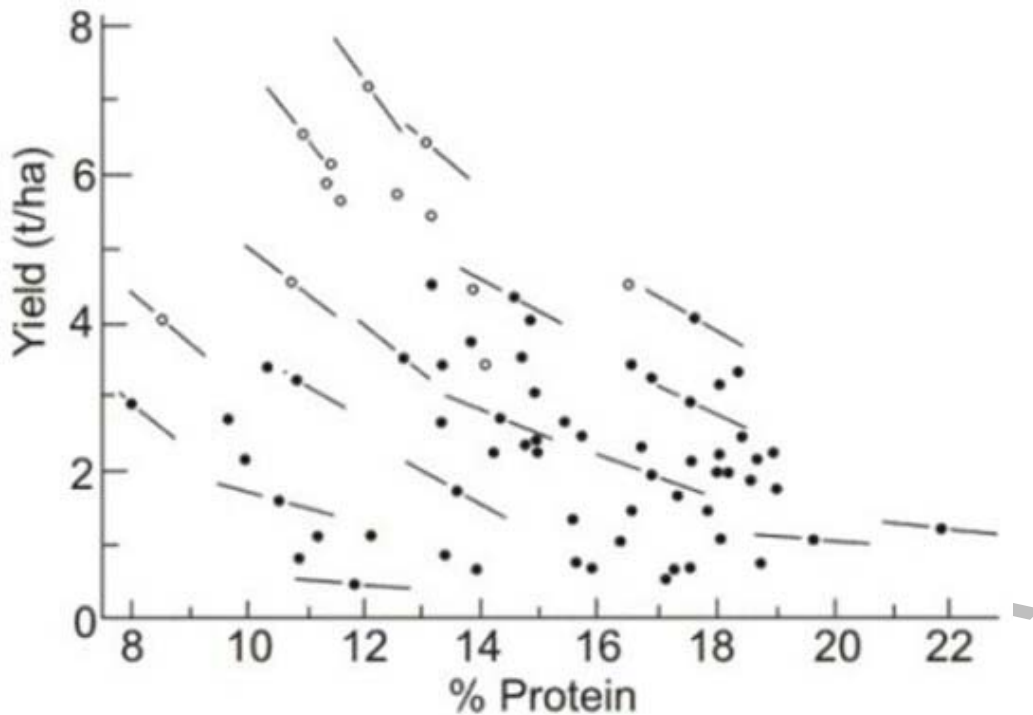


Figure 5. Scatter diagram of yield ( $y$ ,  $t\ ha^{-1}$ ) on protein percent ( $p$ ) for wheats in North America; and Europe. A sample of linear regressions of  $y$  on  $p$  within experiments is also plotted. Taken from Simmonds, N.W. (1995) The relation between yield and protein in cereal grain. *Journal of the Science of Food and Agriculture* **67**, 309-315, with permission of the publishers.

Wild lines of *T. turgidum* var *dicoccoides* (wild emmer) originating from Israel have grain protein contents of up to almost 30 percent and can be used to increase the grain protein content of cultivated durum wheat and bread wheat to up to 20 percent dry wt. However, in this material, as in other wheats, grain protein concentration is negatively correlated with yield components, such as kernel weight and number of kernels per spike.

The genetic control of grain protein concentration is complex, as would be expected for a character determined by a combination of physiological attributes including direct uptake of nitrogen, transport of nitrogen into the grain and remobilization from vegetative tissue. Thus, analyses of the high protein character derived from wild emmer in a durum wheat background have identified genes on a number of chromosomes (1A, 2A, 3A, 5B, 6B), which affect either nitrogen accumulation or partitioning. The development of marker-assisted selection provides a powerful new tool to breed for high protein content, as for other complex multi-genic characters.

### 3. Improving Grain Amino Acid Composition

Comparison of the amino acid compositions of different genotypes of wheat, barley or rye shows little or no variation in the contents of lysine, threonine or other nutritionally essential amino acids, apart from that associated with variation in total grain protein as discussed above. In addition, such variation does not appear to be present in the wild

species with which these cereals can be crossed. Consequently, two approaches have been used to increase the level of variation in these amino acids: mutagenesis and genetic engineering.

### 3.1. High Lysine Mutants

The discovery in 1964 of the *opaque2* high lysine gene of maize led to the search for similar high lysine phenotypes in maize and other cereals. *opaque2* and other mutants subsequently identified in maize (*floury2* and *3*, *opaque6* and *7*, *brittle1* and *2*, *shrunkn1* and *4* and *sugary1*) were all spontaneous mutants identified visually on the basis of changes in storage carbohydrates which also occur (see *Molecular Genetic Improvement of Protein Quality in Maize*). In contrast, only one spontaneous mutant was identified in barley, which was an Ethiopian line (Hiproly) identified by a painstaking screening of lines from the World Barley Collection using combined analyses for dye binding capacity (a measure of total basic amino acids including lysine) and Kjeldahl nitrogen. Subsequently, other workers used a range of chemical and physical mutagens to induce further high lysine mutants with considerable success being achieved by workers at the Risø National Laboratory, Roskilde, Denmark.

Similar high lysine mutations have not been reported in wheat or rye and there has not, to the best of my knowledge, been any systematic attempt to identify spontaneous or induced mutations in these species. In addition, the hexaploid nature of bread wheat means that dominant mutants are more readily identified than recessive mutations, whereas most of the high lysine mutations identified in other cereals are recessive.

Only one of the mutant high lysine barley genes maps to a hordein structural locus (Risø mutant 56, see below) while all of the others map elsewhere on the genome and are loosely termed “regulatory”. Their impact on grain lysine varies from increases of only a few percent compared with the amounts present in isogenic control lines to increases of over 30 percent, and these increases are usually associated with corresponding decreases in the proportions of the lysine-poor hordeins. Although there was initially considerable optimism about the prospects for incorporating some of the high lysine genes into high yielding cultivars for commercial production, this has not proved to be possible. This is because the high lysine phenotype is invariably associated with decreased synthesis of starch, and hence reduced yield. Nevertheless, high lysine barley has proved to be a fascinating topic for study, with three mutants being of particular interest.

The high lysine phenotype of Hiproly is determined by a single recessive gene (*lys*) located on barley chromosome 5H. The original line has about 30 percent more lysine than normal cultivars, but only about 30 percent of the yield. Despite intensive breeding efforts, it has not proved possible to produce commercially viable lines, and this may be due to pleiotropic effects of the *lys* gene on starch accumulation. In contrast to some other high lysine barleys, the amount of hordein is only decreased by about 10-20 percent with no apparent effects on the hordein polypeptide composition. In fact, about half of the increased lysine content of Hiproly results from specific increases in four proteins which together account for about 17 percent of the salt-soluble proteins compared with 7 percent in normal cultivars. These proteins are  $\beta$ -amylase ( $\cong$  5g

percent lysine), protein Z (now known to be a serpin proteinase inhibitor) (7.1g percent lysine) and chymotrypsin inhibitors CI-1 (9.5g percent lysine) and CI-2 (11.5g percent lysine) (Table 4). CI-1 and 2 will be discussed in more detail below.

	Protein mg g <sup>-1</sup>				
	Total Protein	β-amylase	Protein Z	CI-1	CI-2
Mona	110	0.98	2.2	0.24	0.08
Hiproly x Mona <sup>5</sup>	127	4.3	4.3	1.8	0.67
Bomi	98	0.85	2.5	0.23	0.06
Hiproly x Bomi <sup>2</sup>	113	3.2	4.5	1.7	0.38

Based on data of Hejgaard and Boisen (1980) High lysine proteins in Hiproly barley breeding: Identification, nutritional significance and new screening methods. *Hereditas* **93**, 311-320.

Table 4. β-Amylase, protein Z and chymotrypsin inhibitors in normal and high lysine lines derived from Hiproly.

Risø 1508 was induced by treatment of barley with the chemical mutagen ethylinimine. The recessive *lys 3a* gene is also present on chromosome 5H but is not linked to *lys*. Risø 1508 was initially of great interest because of its massive increase in lysine (+36 percent) combined with modest decreases in yield (-23 percent) and 1000 seed weight (-12 percent). However, it has again proved impossible to completely eliminate the yield penalty and the major interest in Risø 1508 has related to its mechanism of action and effects on grain proteins, notably hordeins which are reduced in amount to about a third of that present in the parental line.

Finally, Risø 56 is a γ-ray mutant and is of interest in that the mutant gene (called *Hor2ca*) is actually a deletion of the *Hor2* locus, which encodes the B hordeins. The loss of this major group of hordeins has resulted in compensatory increases in C hordein and in other more lysine-rich proteins, leading to the high lysine phenotype.

Thus, despite a considerable investment in mutation breeding for high lysine barleys, particularly during the 1970s and 1980s, the final results have been negative, due to an inability to separate the high lysine character from the negative effects on grain yield.

### 3.2. Increasing Free Amino Acids

Lysine, threonine and methionine are synthesised by a pathway starting with aspartate, as shown in Figure 6. Detailed studies at the genetic, biochemical and molecular levels have shown that this pathway is feedback regulated, with both positive and negative controls on the first enzyme (aspartate kinase, AK) and on the enzymes controlling the branch points leading to threonine and methionine (homoserine dehydrogenase, HSD) and lysine (dihydrodipicolinate synthase, DHPS).

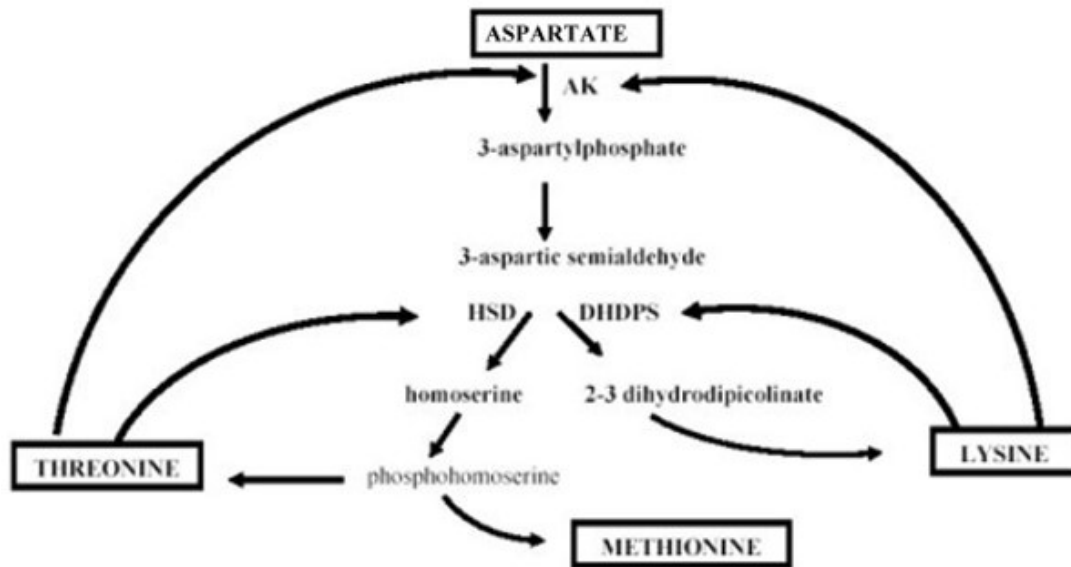


Figure 6. The pathway of synthesis of amino acids derived from aspartic acid. Pathway steps are indicated by arrows. Regulatory feedback steps are shown by loops.

In barley, aspartate kinase occurs as three isoenzymes, which are inhibited by lysine (AKII and AKIII) or threonine (AKI). In addition, the sensitivity to lysine is increased in the presence of S-adenosyl methionine. Thus, the flow into the pathway is regulated by the combined levels of the three major end products: lysine, threonine and S-adenosylmethionine.

HSD also usually occurs in at least two forms, one of which is sensitive to feedback inhibition by lysine. In addition, in most, if not all plant species, these activities are present with AK as a single protein. Similarly, DHPS is highly sensitive to feedback inhibition by lysine, with at least two forms present in wheat. The reader is referred to the Bibliography for further details of these enzymes and of other regulatory aspects of the aspartate pathway that are not relevant to the present discussion.

The presence of multiple isoenzymes and the demonstration of feedback regulation of the enzymes controlling the flux into the aspartate pathway and the two major branch points, led researchers to devise a strategy for overproduction of lysine and threonine. Mutants were induced in barley by treatment with sodium azide and then selected by culturing embryos on a medium containing lysine and threonine. In normal barley embryos, this combination of amino acids should lead to death due to starvation of methionine. Four resistant mutants were isolated, all of which contained increased amounts of free lysine and threonine in their mature grain (Table 5). The increases in threonine were up to 70-fold (in R2506) and of lysine up to 15 to 20-fold. In addition, soluble methionine was increased by about three and six-fold in two mutants (R2506 and R2501, respectively). Detailed biochemical studies of three of the mutants (R2501, R3004 and R3202) showed effects on the regulatory properties of the lysine-sensitive AK isoenzymes (AKII and AKIII).

Plant	Threonine	Lysine	Methionine
Bomi	118	84	24
R2501	6129	1266	149
R2506	9032	1620	72
Bomi	127	79	19
R3004	2376	107	23
R3002	125	109	19
R3004 x R3202 (double)	689	129	21

Amino acid content in nmol g<sup>-1</sup> dry weight.

Taken from Lea, P.J., Blackwell, R.D. and Azevedo, R.A. (1992) Analysis of barley metabolism using mutant genes, in *Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology* (ed. P.R. Shewry), pp181-208, CAB International, Wallington.

Table 5. Soluble threonine, lysine and methionine content in mature grains from wild-type, single and double mutants of barley resistant to lysine plus threonine.

The use of mutagenesis has not had any long-term impact in terms of improved nutritional quality, but has established the scientific basis for the use of transformation, but with feedback insensitive enzymes from bacteria rather than plants. Thus, the expression of feedback insensitive forms of AK and DHPS from the bacterium *E.coli* in tobacco leaves resulted in increases of 14-fold and 15-fold in free threonine and free lysine, respectively, while seed-specific expression of the AK alone gave a seven-fold increase in free threonine and a three-fold increase in free methionine. However, when expressed in barley, the DHPS gene gave only a two-fold increase in free lysine in the grain, while no change was observed with the AK gene. This contrasts with the expression of lysine feedback-insensitive DHPS or DHPS and AK enzymes from bacteria (*Corynebacterium* and *E.coli*) in seeds of maize and canola. Expression of DHPS alone in canola resulted in a 100-fold increase in free lysine and a 2-fold increase in total lysine, while expression of DHPS and AK in soybean resulted in increases of several hundred-fold and up to five-fold in free and total seed lysine, respectively. So far, such increases have not been achieved in cereals, but there appears to be no *a priori* reason why this would not be possible.

### 3.3. Transformation with Genes for High Quality Proteins

Although it may be possible to improve grain quality by increasing free amino acids as discussed above, this approach suffers from the drawback that the pools of these amino acids in normal grain are low, so that massive increases are required to have an impact on the composition of the whole grain. We do not yet know whether such increases can be tolerated by the grain without adverse effects on other aspects of metabolism or development. An alternative approach is to transform the cereals to express high levels of proteins rich in nutritionally-limiting amino acids in the grain.



This has not yet been shown experimentally in barley, wheat or rye, but the feasibility has been demonstrated by work on other species. Most notably, a novel protein containing 31 percent lysine and 20 percent methionine has been designed and a corresponding synthetic gene expressed in transgenic tobacco under the control of the seed-specific soybean  $\beta$ -conglycinin and bean phaseolin promoters. Increases in total seed lysine and methionine occurred which were consistent with the novel protein accounting for up to 2 percent of the total seed proteins. For example, in one homozygous line, the contents of methionine and lysine were increased from 1.53 to 1.83 percent and from 2.41 to 2.88 percent, respectively. Similarly, the expression of a methionine-rich 2S albumin protein from Brazil nut has resulted in increased total seed methionine in tobacco, oilseed rape (by up to 30 percent), and *Vicia narbonensis* (by up to three-fold), while a methionine-rich albumin from sunflower has been used to increase the proportion of methionine in lupin seeds by 94 percent. However, in the latter case this increase was at the cost of cysteine (reduced by 12 percent) and free oxidised sulphur (i.e., sulphate) and there was no effect on the total sulphur content of the seed.

This latter observation suggests that the total supply of sulphur may limit our ability to increase total seed cysteine + methionine. However, in the case of lysine and threonine, there is no evidence that their rate of synthesis would limit the accumulation of high quality proteins. Furthermore, it would be of interest to combine lines engineered for accumulation of free lysine and threonine with genes for high value proteins, in order to determine whether any synergy occurred (for example, by stimulatory effects of free amino acid levels on storage protein gene expression).

Although an *ab initio* designed high quality protein has been successfully expressed in transformed tobacco, it is probably more feasible to transform with genes for naturally occurring high lysine proteins. The best candidate proteins available at the moment are the barley chymotrypsin inhibitors discussed above, particularly CI-2. The major form of CI-2 present in the mature grain, called CI-2A, consists of 84 residues with an  $M_r$  of 9380. It has seven lysine residues and contains no cysteines, and hence no disulphide bonds. The three-dimensional structure (Figure 7) shows a wedge-shaped disc with a single  $\alpha$ -helix and four strands of  $\beta$ -sheet with a left-handed twist. The reactive site (Met<sup>59</sup>-Glu<sup>60</sup>) is present in a loop region (Gly<sup>54</sup>-Tyr<sup>61</sup>) and it is possible to engineer this loop to insert alternative or additional amino acid residues. For example, it has been shown that 10 glutamine residues can be inserted into the loop at Met<sup>59</sup> or used to replace Gly<sup>54</sup>-Tyr<sup>61</sup> without affecting the ability of protein to be expressed in *E.coli* and folded. Similarly, workers at Pioneer Hi-bred (Iowa) have recently patented the engineering of CI-2 to contain additional amounts of lysine and other nutritionally essential amino acids and the expression of the mutant proteins in transgenic plants.

CI-2 has the advantage for use in transformation that substantial increases in amount are tolerated in Hiproly and derived high lysine lines. However, other wild type or engineered high lysine proteins could also be used. Similarly, increases in other nutritionally essential amino acids, such as histidine, tryptophan and threonine, could be achieved by expression of  $\gamma$ -zein from maize ( $\cong$  8 mol per cent His), puroindoline a from wheat (5 Trp out of 120 residues) or D hordein from barley ( $\cong$  8 mol percent Thr), respectively, or by expression of engineered or *ab initio*-designed proteins. The  $\delta$ -zeins

and  $\beta$ -zeins are also rich in methionine ( $\cong$  22 and 11 mol percent, respectively), but the potential for increasing the level of this amino acid may be limited by sulphur availability, as discussed above.

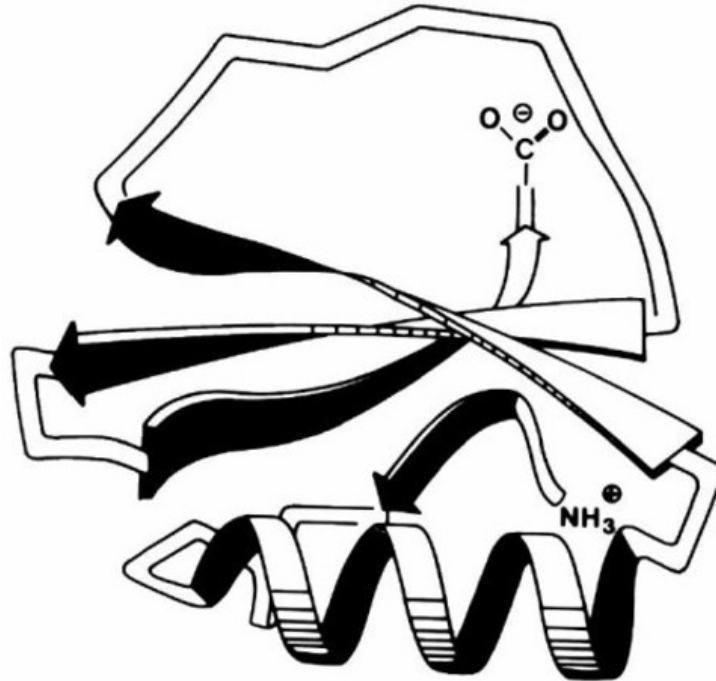


Figure 7. The three-dimensional structure of chymotrypsin inhibitor CI-2. Large arrows denote  $\beta$ -strands; wide ribbons are  $\alpha$ -helix; and narrow ribbons are turns or unclassified structures. The small arrow indicates the reactive (inhibitory) site. Adapted from McPhalen, C.A., Svendsen, I., Jonassen, I. and James, M.N.G. (1985) Crystal and molecular structure of chymotrypsin inhibitor 2 from barley seeds in complex with subtilisin Novo. *Proceedings of the National Academy of Science, USA* **82**, 7242-7246, with permission of the authors.

#### 4. Conclusion

Our ability to increase the amount of total protein in the grain of wheat, barley and rye is currently limited by several factors. Firstly, there is limited genetic variation available in this character and no major genes have been identified. Secondly, we currently know very little about the mechanisms that control grain protein accumulation at the molecular, biochemical and physiological levels, which limits our ability to develop strategies based on genetic engineering. In contrast, there are clear opportunities to improve the amino acid composition of the grain to increase the proportions of nutritionally essential amino acids, based on genetic engineering to increase the amounts of free amino acids and/or specific high quality proteins, the latter including mutated and *ab initio* designed proteins, as well as naturally occurring proteins from cereals or other species.

However, any attempts to manipulate the protein composition of grain must also be considered, in relation to their impact on the processing properties. This is particularly

important for wheat where the grain storage proteins form the gluten fraction, which determines the functional properties for making bread and other foods. It is essential that these properties should be maintained or even enhanced if the nutritionally-improved lines are to be culturally acceptable.

## Acknowledgements

Long Ashton Research Station receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

## Glossary

<b>Albumin:</b>	A protein soluble in water.
<b>Amino acids:</b>	The building blocks of proteins. Some 20 different amino acids occur widely in proteins.
<b>β-amylase:</b>	An enzyme that contributes to the breakdown of starch.
<b>Cereals:</b>	Grass species that are cultivated for their seeds. The three “major” cereals are wheat, maize and rice, with barley, oats, sorghum, rye and millets being less widely cultivated.
<b>Chymotrypsin inhibitors:</b>	Proteins that inhibit the activity of the proteolytic enzyme chymotrypsin.
<b>β-conglycinin:</b>	A globulin storage protein of soybean seeds.
<b>Disulphide bonds:</b>	Cross-links formed between two residues of the amino acid cysteine, present within a protein or in separate proteins.
<b>Essential amino acids:</b>	Amino acids that cannot be synthesized by animals and therefore must be provided in their diet. They include lysine, threonine (deficient in most cereal seeds), cysteine and methionine (sulphur-containing amino acids which are deficient in most legume seeds) and tryptophan.
<b>Functional properties:</b>	Properties of plant or animal components that contribute to their behaviour in food processing. For example, proteins may exhibit a range of properties including foaming, gelation, visco-elasticity and emulsification.
<b>Gliadins:</b>	Wheat grain prolamins that are readily soluble in alcohol-water mixtures.
<b>Globulin:</b>	A protein soluble in dilute salt solution (e.g., 0.5-1.0M NaCl).
<b>Glutelin:</b>	A plant protein that is soluble only in the presence of dilute acid or alkali, chaotropic agents (e.g., urea) or detergents.
<b>Gluten:</b>	The prolamins (gliadin + glutenin) proteins of wheat grain, which form a visco-elastic network in dough.
<b>Glutenins:</b>	The alcohol-insoluble (polymeric) prolamins of wheat.
<b>Hordeins:</b>	The prolamins of barley grain.
<b>Inbreeding:</b>	Plants that are self-pollinated.
<b>Outbreeding:</b>	Plants that are cross-pollinated.
<b>Ploidy:</b>	The number of chromosome sets per cell. Most plants, including some wheat species, are <b>diploid</b> with two chromosome sets, but polyploid species (with more than two chromosome sets) also occur. Thus bread wheat is <b>hexaploid</b>

(six chromosome sets) while durum (pasta) wheat is **tetraploid** (four chromosome sets).

**Prolamins:**

A type of protein present only in cereal grains. Prolamins are classically defined on solubility in aqueous alcohol (60-70% (v/v) ethanol), but the term is now taken to include related proteins, which are present as alcohol-insoluble polymers.

**Phaseolin:**

A globulin storage protein from bean (*Phaseolus* spp) seeds.

**Puroindolines:**

A type of tryptophan-rich protein present in wheat grains.

**Non-ruminants:**

Animals that have simple stomachs and therefore need to be provided with essential amino acids in their diet.

**Ruminants:**

Animals that have an additional stomach (rumen) that contains bacteria, which are able to synthesise essential amino acids from other compounds. These animals therefore do not need to be provided with essential amino acids in their diet.

**Secalins**

The prolamin storage proteins of rye grains.

**Zeins:**

The prolamin storage proteins of maize grains.

**Bibliography**

Bright, S.W.J. and Shewry, P.R. (1983) Improvement of protein quality in cereals. *CRC Critical Reviews in Plant Sciences* **1**, 49-93. [A review of strategies to improve quality, based on high lysine mutants and increasing free amino acids].

Falco, S.C., Guida, T., Locke, M., Mauvais, J., Sanders, C., Ward, R.T. and Webber, P. (1995) Transgenic canola and soybean seeds with increased lysine. *Biotechnology* **13**, 577-582. [This demonstrates the use of transformation with bacterial enzymes to increase free lysine and threonine].

Galili, G. (1995) Regulation of lysine and threonine synthesis. *The Plant Cell* **7**, 899-906. [A review of lysine and threonine biosynthesis, and how this can be manipulated].

Keeler, S.J., Maloney, C.L., Webber, P.Y., Patterson, C., Hirata, L.T., Falco, S.C. and Rice, J.A. (1997) Expression of *de novo* high-lysine  $\alpha$ -helical coiled-coil proteins may significantly increase the accumulated levels of lysine in mature seeds of transgenic tobacco plants. *Plant Molecular Biology* **34**, 15-29. [A novel approach to designing and expressing high quality proteins].

Lea, P.J., Blackwell, R.D. and Azevedo, R.A. (1992) Analysis of barley metabolism using mutant genes. *Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology* (ed. P.R. Shewry), 181-208. CAB International, Wallington. [This includes a review of progress in manipulating amino acid biosynthesis by mutagenesis in barley].

Molvig, L., Tabe, L.M., Eggum, B.O., Moore, A.E., Craig, S., Spencer, D. and Higgins, T.J.V. (1997) Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (*Lupinus angustifolius* L.) expressing a sunflower seed albumin gene. *Proceedings of the National Academy of Science USA* **94**, 8393-8398. [This demonstrates the potential for increasing seed methionine by expression of a methionine-rich protein].

Munck, L., Karlsson, K.-E., Hagburg, A. and Eggum, B.O. (1970) Gene for improved nutritional value in barley seed protein. *Science* **168**, 985-987. [The classic paper describing the identification of the first high lysine barley mutant, Hiproly].

Shewry, P.R. and Casey, R. (Eds.) (1999) *Seed Proteins*. Kluwer Academic Publishers, Dordrecht. [This monograph contains detailed accounts of cereal seed proteins].

Shewry, P.R., Williamson, M.S. and Kreis, M. (1987). Effects of Mutant Genes on the synthesis of storage components in Developing Barley Endosperms. In *Mutant Genes That Affect Plant Development*. (Ed. H. Thomas and D. Grierson). Cambridge University Press, Cambridge, pp. 95-118. [A review of mutant high lysine genes in barley].

Simmonds, N.W. (1995) The relation between yield and protein in cereal grain, *Journal of the Science of Food and Agriculture* **67**, 309-315. [A review of the relationship between grain protein content and yield in cereals].

### **Biographical Sketch**

**Peter Shewry** is an authority on the structures and properties of plant proteins, and in particular the seed storage proteins of cereals. He has published extensively on the proteins of barley, wheat and rye grains over a period of 25 years and contributed to many collaborative research programmes and to international conferences and workshops. In 2000 he was awarded the Thomas Burr Osborne medal by the American Association of Cereal Chemists. He is currently Associate Director and Head of the Crop Performance and Improvement Division at Rothamsted Research. He is also Professor of Agricultural Sciences in the University of Bristol.

### **To cite this chapter**

**P. R. Shewry.**, (2004/Rev.2006), IMPROVING THE PROTEIN CONTENT AND QUALITY OF TEMPERATE CEREALS: WHEAT, BARLEY AND RYE, in *Impacts of Agriculture on Human Health and Nutrition*, [Eds. Ross M. Welch, and Ismail Çakmak], in *Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [<http://www.eolss.net>]