# An integrated approach to predicting end - product quality

**Békés F and Morell M**
*CSIRO Food Futures National Research Flagship, Canberra ACT 2601*

## INTRODUCTION

An overview is given in this paper on the recent achievements of combining the *in vitro* and *in vivo* methodologies of cereal science with molecular techniques and mathematical modelling in our laboratory. The paper introduces a novel prediction procedure relating the individual and interactive contribution of HMW and LMW glutenin alleles to specific dough parameters and end-product quality attributes as well as single parameters which seems to be capable of providing valuable information on the end-product specific quality of a sample

## RESULTS AND DISCUSSION

### Basic methodologies relating quality attributes to protein composition

Most of our knowledge about the 'genetics of quality' or about the relationships between chemical composition and quality attributes derives from two different experimental approaches: i) direct measurements of quality traits on samples with systematically altered chemical composition (Bekes and Gras, 1999); ii) relating quality and chemical composition/genetics of large sample populations using statistical methods. In both cases, the complexity of relating composition to quality is derived from the fact that the question can (and has to be) investigated on different levels of protein composition. There are certain relationships of quality with protein content, and/or with the ratio of polymeric and monomer proteins, and/or width the ratio of HMW and LMW glutenin subunits etc. The presence and the relative levels of individual polypeptides or the size distribution of polymeric glutenin fraction also can be related to quality attributes. The G, E and G x E effects alter each of these levels simultaneously, so complex methodologies are required to separate and evaluate the effects on quality at the different structural levels.

### *In vitro* methodologies

The concept of *in vitro* direct methodologies is based on the reconstitution technique (MacRitchie 1987): separating flour components and putting them back without changing the functionality of the isolated components. This allows replacing a component with those isolated from another samples and/or modifying the relative amount in the reconstituted sample, so chemical composition can be altered both qualitatively and quantitatively. The capabilities of *in vitro* methodologies have been multiplied by the development of so-called incorporation procedure which allows building the supplemented glutenin subunits into the polymeric glutenin structure through a partial reduction/re-oxidation process (Bekes et al 1994a).

Two basic *in vitro* methodologies have been developed and successfully applied in the recent years in our laboratory: i) the 'base flour mehod' where the chemical composition of a base flour is altered systematically by supplemented (simply added or incorporated) compounds or ii) the method where a 'model dough' is formed by putting isolated flour components (starch, soluble proteins, lipids, gliadins and in vitro polymerised glutenin subunits together

### Indirect statistical methods to relate genetic/chemical data to quality

It is well known that wheat varieties differ in their bread-making ability and that the endosperm proteins have a major influence on bread making quality. Wheat varieties at the same protein level were found to differ in their bread making quality, giving the first indication that protein quality, as well as the amount is important for good bread-baking quality. The statistical approach, relating the protein composition to certain quality traits by statistical means is the most frequently used methodology to relate structure/composition to function/functionality in cereal science. Water absorption is predicted as a function of protein content, the amounts of pentosans and the level of starch damage in the sample; the amount of un-extractable polymeric proteins is a good measure of dough strength (Gupta et al 1993), extensibility can successfully be related to the amount of polymeric protein in the flour (glutenin% multiplied by the protein content of the flour). The most frequently used tool relating protein composition to quality is the Payne score (1987) providing a single number to estimate dough strength from the HMW glutein allelic composition. There are several attempts to also involve the LMW glutenin alleles in similar mathematical formulas (Gupta et al,1991; Eagles et al 2002). The method of Eagles et al (2002), applying a sophisticated statistical approach of analysis of variance, is capable to describe the effects of both HMW and LMW GS individually and the pair-wise interactions among the alleles.

In recent years the molecular/genetic studies on wheat have provided a new approach for examining the regions of chromosomes affecting significant variation in flour processing properties. The determination of the genetic loci that underpin the observed variation in processing traits is possible as a result of the extensive genetic maps that are now available. Through the generation of doubled haploid lines, from a cross of interest, material can be readily grown in different environments, and analyzed by different laboratories. Advances in the genetic mapping of wheat, the molecular interpretation of flour processing traits and large–scale sequencing of genes expressed in endosperm tissue are currently converging to define the genes that under-pin key quality traits. (Appels *et al.* 2000). QTL analysis of quality traits of wheat through the brief introduction of molecular genetics, cereal chemistry and statistical methods developed and applied recently in this area (Appels et al, 2001, Bekes et al, 2002)

Parallel investigations in relation to quality traits on gene level such as QTL analysis and on gene product level such as protein composition are complementing each other (Bekes et al 2001, Cornish et al, 2001). While QTL analysis is a successful tool for discovery for quality related genes, statistical analysis on large sample populations with known genetic background such as DH populations provide essential information on the contribution of certain storage protein alleles to certain quality traits (Appels et al, 2001). Application of Multi-parent Advanced Generation Inter-cross (MAGIC) populations derived from breeder-relevant germplasm as a platform for a new generation of gene-trait analysis in crop species, multiply the information and application of genetic populations. In such populations the greater diversity of multiple alleles and large sizes allow modelling of epistatic effects for multiple alleles greater recombination and an ability to reduce the genomic interval down to a more manageable size than previously possible whilst still sampling breeder relevant germplasm (Cavanagh at al 2008).

Data derived from research programs such as National Wheat Molecular Marker Program (NWMMP) established in 1996 in Australia (Appels et al, 2001) or the National Wheat Quality Evaluation Program provide an excellent basis for relating genetic and chemical information to quality attributes on a large population of Australian wheat samples. Data-mining in these sets of data – and in similar industry driven results - can serve as the basis of a new level of knowledge about the relationships between protein composition and quality attributes.

## Relating the glutenin allelic composition to quality

A database consisting of more than 3000 samples of Australian germplasm with detailed protein composition as well as quality data has been used i) to develop approaches to predict quality traits from genetic and chemical measurements; ii) to better understand the role and contribution of different genes and alleles in determining quality; iii) to define questions that can be further answered through genetic analysis and iv) to provide tools for breeders. All data utilized from NWMMP and NWQEP the programs have been anonymously coded so that linkages between specific lines and outputs can not be retrieved, only linkages between genes and outputs**.** Quality data, used were treated to decrease the effect of growing conditions by normalizing them to those of Janz appearing in each of the subsets used in this study.

The data set used for analysis represent a large range of biodiversity for the HMW and LMW glutenin coding genes, 17 (3+10+4) HMW and 16 (6+5+5) LMW glutenin alleles have been involved in the analysis (appearing minimum in three individual samples). Numbers in brackets show the number of alleles of A, B and D genomes for the Glu1 and Glu3 genes, respectively. Investigation of the distribution of these alleles in the combined sample population showed that only the 3.45% of the theoretically possible number of combinations (18000) appeared in the population indicating that extensive biodiversity exists in Australian wheat germplasm. This biodiversity has not been fully utilised to date, to improve quality.

Developing an incident matrix from the allelic composition of the samples, a complex multiple regression type of statistical analysis has been carried out providing factors to describe the contribution of each of the HMW and LMW glutenin alleles and their pairwise interactions to dough strength and extensibility (Figure 1.) Despite the completely different mathematical model used, the resulting parameters show a remarkably good agreement to those of Eagles et al (2002), and in the ranking of HMW alleles to the original Payne score. This agreement indicates that the approach of relating allelic composition to quality attributes is possible with careful data selection and applying robust mathematical tools. The genetic potential of a line, with a certain combination of alleles on the six glutenin coding loci, can be meaningfully predicted where both the contribution of the individual alleles and their pair-wise interactions play equally important role

Because of the large contribution of allele-allele interactions (see Figure 2 as example), the different allelic combinations, rather than the individual glutenin alleles should be targeted in breeding situations to develop new lines with certain quality attributes. In the light of this, the real value of a certain allele cannot be judged 'simply' by 'putting the allele in a common background'. A set of backgrounds has to be applied to be able to realize the interaction-potential of the allele of interest.

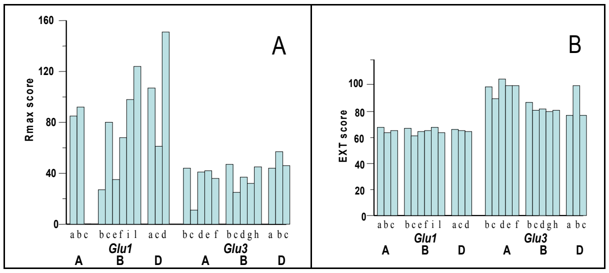

**Figure 1** The individual contribution of HMW and LMW glutenin subunit alleles to dough strength (A) and extensibility (B).

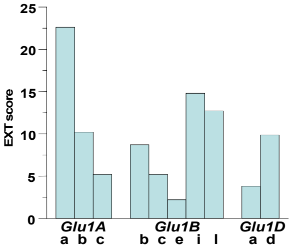

**Figure 2** Comparison of the effects on extensibility related to the interaction HMW-GS alleles with the *GLU-3Bb* allele

From the point of view of its relationship to quality, both the qualitative and quantitative aspects of protein composition of wheat are equally important. Therefore both the genetic makeup (allelic composition) and the total and relative expression levels of the storage protein genes are essential factors determining flour quality. The structure of our incident matrix-based model is capable to building in quantitative factors (such as protein content, glutenin to gliadin ratio etc) into the predictive model resulting in a relatively simple scoring system to predict the actual quality of a sample (Figure 3).

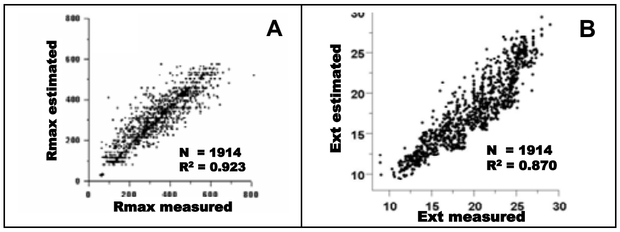

**Figure 3**. Comparing measured and predicted dough strength (A) and extensibility (B) results using allelic composition and protein content data.

## Introducing Protein Quality Index

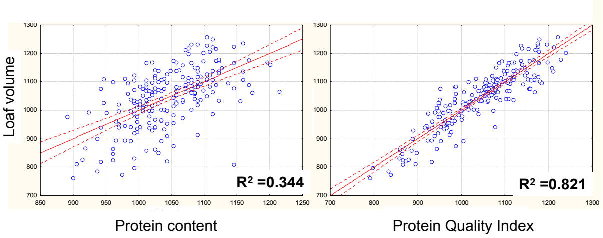

**Figure 4**. Predictive models for loaf volume using protein content and Protein Quality Index

Both protein content and protein composition are extremely important determining flour quality (dough properties). Thus, protein content of a sample alone does not describe its quality. Based on our work on a large population of Australian wheat samples, it is possible to develop single parameters (Protein quality Index, PQI), which can serve as an end-product-specific measure of quality and therefore could be the basis of a future wheat classification system replacing the existing - often misleading - protein content-based system.(Figure 4) The mathematical formula of PQI for bread-making quality was constructed to be able to describe the need for balanced dough strength and extensibility using the Morup-Olesen algorithm and is capable to provide a different scale for pan bread and sponge & dough bread quality.

Protein Quality Index can be used to characterise a different growing locations and select optimal genetic make-up for storage protein genes significantly better than protein content.

## REFERENCES

Appels, R., Gustafson, J.P., O'Brien, L., 2001. Wheat breeding in the new century: applying molecular genetic analysis of key quality and agronomic traits. Aust. J. Agric. Res. 52**,** 1043-1417.

Békés, F., and Gras, P.W. 1999 In vitro studies on gluten protein functionality. Cereal Foods World, 44. 580-586.

Békés, F., Gras, P.W., Anderssen, R.S., Appels, R., 2001. Quality traits of wheat determined by small-scale dough testing methods. Aust. J. Agric. Res. 52, 1325-1338.

Békés, F., Gras, P.W., and Gupta, R.B. 1994 Mixing properties as a measure of reversible reduction/ oxidation of doughs. Cereal Chem. 71. 44-50.

Békés, F., Kemény, S., and Morel, M. 2006 An Integrated approach to predicting end-product quality of wheat. European Journal of Agronomy. 25: 155-162

Békés, F., Lukow, O., Uthayakumaran, S., Mann, G., 2003. Small-scale dough testing methods. In: Shewry, P.R., Lookhart G. (Eds), Wheat Gluten Protein Analysis. AACC, St Paul, USA, pp. 173-198 .

Békés, F., Ma, W., Gale, K., 2002. QTL analysis of wheat quality traits. Acta Agron. Hung. 50, 249-262.

Cavanagh, C., Morell, Mackay, I. and Powell, W. 2008 From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Current Opinion in Plant Biology, 11. 215–221.

Cornish, G.B., Békés, F. Eagles, H.A. and Payne, P.I. 2006 Chapter 8. Prediction of Dough Properties for Bread Wheats. In: Gliadin and glutenin. The unique balance of wheat quality., Eds.: Wrigley, C.W., Békés, F., and Bushuk, W., Pp 243-280. AACCI Press, St Paul, Min., USA

Cornish, G.B., Bekes, F., Martin, D.J., Allen H., 2001. Seed storage proteins linked to quality traits in Australian wheat crosses. Aust. J. Agric. Res. 52, 1339-1348.

Gupta, R. B., Khan, K., and MacRitchie, F. 1993. Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein. J. Cereal Sci. 18:23-41.

Gupta, R. B., Paul, J. G., Cornish, G. B., Palmer, G. A., Békés, F., and Rathjen, A. J. 1994. Allelic variation at glutenin subunit and gliadin loci, *Glu-1*, *Glu-3* and *Gli-1*, of common wheats. I. Its additive and interaction effects on dough properties. J. Cereal Sci. 19:9-17.

Eagles, H.A., Hollamby, G.J., Gororo, N.N., Eastwood, R.F., 2002. Estimation and utilisation of glutenin gene effects from the analysis of unbalanced data from wheat breeding programs. Aust. J. Agric. Res. 53, 1047-57.

Eagles, H.A., Eastwood, R.F., Hollamby, G.J., Martin, E.M., Cornish, G.B., 2004. Revision of the estimates of glutenin gene effects at the *Glu-B1* locus from southern Australian wheat breeding programs. Aust. J. Agr. Res. 55, 1093-1096.

Payne, P. I. 1987. Genetics of wheat storage protein and the effect of allelic variation on breadmaking quality. Ann. Rev. Plant Physiol. 38:141-153.