

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**FREQUENCY OF OCCURRENCE OF NOVEL MILK PROTEIN
VARIANTS IN A NEW ZEALAND DAIRY CATTLE STUDY
POPULATION**

A thesis presented in partial fulfilment of the requirements
for the degree of
Master of Science in Biochemistry,
Massey University,
New Zealand

Richard Gordon Burr
August, 1996

ABSTRACT

Since the discovery of genetic polymorphism within milk protein genes, a considerable volume of research has been published relating milk protein genetic variants and milk production properties. Polymorphism of milk proteins can result in two effects:

(a) changes in the biological and physico-chemical properties of systems containing the variant protein,

(b) changes in the synthesis level of variant proteins.

As a result several studies of milk protein variants have identified phenotypes which may be commercially advantageous for specific products.

Currently employed methods to determine milk protein phenotypes are generally limited to electrophoretic techniques. The gel electrophoretic techniques commonly used are able to detect most milk protein variants that differ by their net electrical charge. However single amino acid substitutions that result in a change in net charge account for only 25% of the possible substitutions that could occur. The remaining 75% of potential variants are the result of a neutral residue substituted by another neutral residue - a 'silent' variant. Thus it is likely that some substitutions, and hence genetic variants have gone undetected in the past.

The purpose of this study was to develop new methods for determining the phenotype of milk proteins, and to determine the frequency of occurrence of silent or other novel variants in a New Zealand dairy cattle study population.

Polyacrylamide gel electrophoresis (PAGE), free zone capillary electrophoresis (CE), peptide mapping by reverse-phase HPLC and electrospray mass spectrometry (ESI-MS) were used in the characterisation of milk proteins purified from 109 individual dairy cows.

Three different PAGE systems were used. Alkaline-urea PAGE enabled the detection of α_{S1} -casein variants B and C, β -casein variants group A (variants A¹, A² and A³) and B, and κ -casein variants A and B in the study population. Beta-casein variants A¹, A² and A³ were subsequently resolved in an acid-urea PAGE system. The whey proteins were very poorly resolved in PAGE systems containing urea. Alpha-lactalbumin A, and β -lactoglobulin

(β -LG) variants A and B were resolved in a non-denaturing 'native' PAGE system. The frequencies of the various milk protein variants corresponded closely to figures previously published.

A free zone CE method that is able to resolve β -LG variants A, B and C was used to check the phenotype of purified β -LG samples. Three samples previously typed as β -LG BB were subsequently determined to be β -LG CC; one sample typed as β -LG BB was re-assigned as β -LG BC. This highlighted the limitations of PAGE systems for the detection of known variants.

Tryptic hydrolysis of purified casein proteins and β -LG, followed by reverse-phase HPLC separation of the resultant peptides was used to create peptide 'maps' of the hydrolysis products. Differences in peptide maps were noted between protein variants. The differences corresponded to peptides containing a substitution site. All samples analysed in this way contained more peptide peaks than expected. Analysis revealed that some were the result of incomplete digestion; others the result of chymotryptic-like cleavages. No aberrant peptide maps, indicative of a silent mutation, were detected.

Purified casein proteins and β -LG were subjected to ESI-MS for mass analysis. The mass of each protein species was determined as follows:

| Protein | Average mass | Std. dev. |
|--------------------------------|--------------|-----------|
| α_{s1} -CN B-8P | 23614.9 Da | 1.2 Da |
| α_{s2} -CN A-11P | 25228.9 Da | 1.5 Da |
| β -CN A ¹ -5P | 24023.9 Da | 3.1 Da |
| β -CN A ² -5P | 23983.5 Da | 1.8 Da |
| β -CN B-5P | 24092.6 Da | n.d. |
| κ -CN A-1P | 19038.8 Da | 1.5 Da |
| κ -CN B-1P | 19003.8 Da | n.d. |
| β -LG A | 18362.6 Da | 1.0 Da |
| β -LG B | 18277.0 Da | 0.9 Da |
| β -LG C | 18287.2 Da | 0.6 Da |

In all cases the experimentally determined mass corresponded to the mass calculated from published primary sequences of milk protein variants.

In addition to the expected β -LG variant in each mass spectrum, additional species were detected differing from the mass of the β -LG species by increments of approximately 324 Da. Although less pronounced, the +324 Da molecular weight species were also detected in a sample of β -LG purchased from the Sigma Chemical Company. The additional species were also detected in whey prepared by ultra-centrifugation, although at a much lower level.

The 324 Da molecular weight adducts observed in ESI-MS spectra of purified β -LG are consistent with an addition of a lactosyl residue to the protein. The observation that these species remain after heat denaturation, reduction and RP-HPLC treatment suggest that the linkage is covalent. Lactulosyl-lysine is known to form in milk products during some processing conditions, particularly during heating. The observation of these glycated species in gently treated, unheated milk suggests that glycation may occur to some extent in the udder of the cow.

The association of the 324 Da molecule with β -LG does not alter the charge, molecular weight or hydrophobicity sufficiently to be detected by PAGE, CE or RP-HPLC.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following:

New Zealand Dairy Research Institute for the opportunity to pursue a higher degree, particularly Dr Mike Boland for his guidance in the direction of this course of study.

Foundation for Research, Science and Technology and the New Zealand Dairy Board for funding.

Drs Chris Moore (senior lecturer, Biochemistry Department, Massey University) and Jeremy Hill (Section Manager, Food Science Section, New Zealand Dairy Research Institute) for their patience, tolerance and supervision during this course of study.

Jo Mudford and Carole Flyger of the Biochemistry Department, Massey University for peptide sequencing.

Professor Brian Nicholson of the Chemistry Department, Waikato University for arranging access to the department's mass spectroscopic facilities.

Wendy Jackson of the Chemistry Department, Waikato University for electrospray - mass spectrometric analysis of peptides and proteins. Particularly for her invaluable help with data analysis and instructions on how to operate the mass spectrometer.

John Allen and Daryl Rowan of the Mass Spectrometry Unit of AgResearch, Palmerston North for providing FAB-MS analysis.

Professor André Ng-Kwai-Hang of the Animal Science Department, McGill University for his encouragement, experience and enthusiasm in establishing the foundations of this work.

All of my colleagues at NZDRI, particularly Christina Coker, Wayne Thresher and Lawrie Creamer for their unflagging support, encouragement and many useful discussions on aspects of this work .

Lastly, and most importantly I wish to thank my family for their tolerance and forgiveness for many missed good night stories.

TABLE OF CONTENTS

| | |
|---|------|
| Abstract | 1 |
| Acknowledgements | v |
| Abbreviations and terminology | xvii |
| | |
| Introduction | 1 |
| | |
| Literature Review | 7 |
| | |
| 1. Milk composition | 7 |
| | |
| 2. Heterogeneity of Bovine Milk Proteins | 8 |
| 2.1 Caseins | 9 |
| 2.1.1 α_{S1} -CN | 9 |
| 2.1.2 α_{S2} -CN | 10 |
| 2.1.3 β -CN | 11 |
| 2.1.4 κ -CN | 12 |
| 2.2 Whey Proteins | 14 |
| 2.2.1 α -La | 14 |
| 2.2.2 β -LG | 15 |
| 2.2.3 Serum albumins and immunoglobulins | 17 |
| | |
| 3. The molecular basis of polymorphism in milk proteins | 19 |
| | |
| 4. Post translational modifications | 21 |
| 4.1 Phosphorylation | 22 |
| 4.2 Glycosylation | 23 |
| 4.3 Disulphide cross-linking | 24 |
| | |
| 5. Genetic organisation of the milk proteins | 26 |
| | |
| 6. Significance of polymorphism | 29 |
| 6.1 Milk production | 29 |

| | | |
|----------|--|----|
| 6.2 | Milk composition | 30 |
| 6.3 | Milk properties | 31 |
| 6.3.1 | Heat stability | 31 |
| 6.3.2 | Association, dissociation and aggregation behaviour of β - LG | 33 |
| 6.3.3. | β -LG interaction with κ -CN | 34 |
| 6.3.4. | Heat induced gelation of β -LG | 34 |
| 6.3.5. | Cheese making | 34 |
| 6.3.5.1. | Renneting | 34 |
| 6.3.5.2. | Syneresis | 36 |
| 6.3.5.3. | Yield and quality | 36 |
| 6.3.6 | Other effects | 37 |
| 7. | Detection of Polymorphism in Milk Proteins | 39 |
| 7.1 | Electrophoresis | 39 |
| 7.1.1 | Paper electrophoresis | 39 |
| 7.1.2 | Starch gel electrophoresis | 40 |
| 7.1.3 | PAGE | 40 |
| 7.2 | Isoelectric focusing (IEF) | 40 |
| 7.3 | 2-D Electrophoresis | 41 |
| 7.4 | Liquid chromatography | 42 |
| 7.4.1 | Ion-exchange | 42 |
| 7.4.2 | Reverse phase | 43 |
| 7.5 | Peptide Mapping | 43 |
| 7.6 | DNA Sequencing | 44 |
| 7.7 | Mass Spectroscopy | 45 |

| | |
|--|----|
| Experimental | 53 |
| 1 Preliminary Sample Preparation | 54 |
| 1.1 Preparation of Skim Milk | 54 |
| 1.2 Casein and Whey Preparation | 54 |
| 2 Determination of Electrophoretic Phenotype | 55 |
| 2.1 Materials and Methods | 55 |
| 2.1.1 Alkaline Urea PAGE | 55 |
| 2.1.2 Acid Urea PAGE | 56 |
| 2.1.3 Native PAGE | 56 |
| 2.2 Results | 57 |
| 2.2.1 Alkaline-urea PAGE | 59 |
| 2.2.2 Acid-urea PAGE | 61 |
| 2.2.3 Native PAGE | 61 |
| 2.3 Discussion | 62 |
| 3 Milk Protein Purification | 65 |
| 3.1 β -lactoglobulin purification | 65 |
| 3.1.1 Materials and Methods | 65 |
| 3.1.2 Results | 65 |
| 3.2 Casein purification | 66 |
| 3.2.1 Materials | 66 |
| 3.2.2 Methods | 67 |
| 3.2.3 Results | 68 |
| 3.2.4 Discussion | 69 |
| 4 β -LG Phenotyping by Capillary Electrophoresis | 71 |
| 4.1 Methods and Materials | 71 |
| 4.2 Results | 72 |
| 4.3 Discussion | 73 |

| | | |
|-------|--|-----|
| 5 | Trypsin Purification | 74 |
| 5.1 | Materials and Methods | 74 |
| 5.2 | Results | 74 |
| 6 | Tryptic Hydrolysis | 76 |
| 6.1 | Methods and Materials | 76 |
| 6.1.1 | Hydrolysis | 76 |
| 6.1.2 | Peptide Mapping | 77 |
| 6.1.3 | Peptide sequencing | 79 |
| 6.1.4 | Fast atom bombardment - mass spectrometry (FAB-MS) | 79 |
| 6.2 | Results | 79 |
| 6.2.1 | α_{S1} -CN | 81 |
| 6.2.2 | α_{S2} -CN | 82 |
| 6.2.3 | β -CN | 85 |
| 6.2.4 | κ -CN | 88 |
| 6.3 | Discussion | 89 |
| 7 | Mass Spectroscopy | 91 |
| 7.1 | Materials and Methods | 91 |
| 7.2 | Mass Spectrum Data Processing | 92 |
| 7.3 | Results | 93 |
| 7.3.1 | α_{S1} -CN | 93 |
| 7.3.2 | α_{S2} -CN | 95 |
| 7.3.3 | β -CN | 96 |
| 7.3.4 | κ -CN | 97 |
| 7.3.5 | Casein | 100 |
| 7.3.6 | β -LG | 100 |
| 7.4 | Discussion | 104 |
| 8 | Analysis of β -LG Adduct Species | 107 |
| 8.1 | Reducing sugar test | 108 |
| 8.1.1 | Methods and materials | 108 |
| 8.1.2 | Results | 109 |
| 8.1.3 | Discussion | 109 |
| 8.2 | Nature of adduct association | 110 |

| | | | |
|-----|-------|--|-----|
| | 8.2.1 | Materials and methods | 110 |
| | 8.2.2 | Results | 111 |
| | 8.2.3 | Discussion | 113 |
| 8.3 | | Preparation of native whey proteins | 115 |
| | 8.3.1 | Methods and materials | 116 |
| | 8.3.2 | Results | 116 |
| | 8.3.3 | Discussion | 117 |
| 9 | | Discussion and conclusion | 119 |
| 10 | | Bibliography | 123 |
| 11 | | Appendix | 163 |
| | 11.1 | Appendix A - Liquid chromatography time and parameter programmes for RP-HPLC separation of milk protein tryptic hydrolyses on Shimadzu system | 163 |
| | 11.2 | Appendix B - FAB-MS spectra of α_{s2} -CN A tryptic peptides. | 164 |
| | 11.3 | Appendix C - ESI-MS data spreadsheets and plots of observed masses for milk proteins analysed | 166 |
| | 11.4 | Appendix D - Statistical analysis of ESI-MS determined masses | 170 |
| 12 | | Publications | 175 |
| | 12.1 | Burr, R., Moore, C.H., Hill, J.P. (1996) Evidence of multiple glycosylation of bovine β -lactoglobulin by electrospray ionisation mass spectrometry | 175 |
| | 12.2 | Burr, R. (1996) Detection of "silent" milk protein variants | 193 |
| | 12.3 | Burr, R., Moore, C.H., Otter, D.O., Hill, J.P. (1996) Use of electrospray ionisation mass spectrometry to elucidate the multiple glycation of bovine β -lactoglobulin | 194 |

FIGURES

| | | |
|---------------|--|----|
| Figure 1 | Distribution of major milk proteins and peptides in bovine milk | 8 |
| Figure 2 | Primary structure of α_{S1} -CN B-8P | 10 |
| Figure 3 | Primary structure of α_{S2} -CN A-11P | 11 |
| Figure 4 | Primary structure of β -CN A ² -5P | 12 |
| Figure 5 | Primary structure of κ -CN B-1P | 13 |
| Figure 6 | Primary sequence of α -Lac B | 15 |
| Figure 7 | Primary sequence of β -LG B | 16 |
| Figure 8 | Primary sequence of bovine serum albumin | 18 |
| Figure 9 | Phylogenetic relationships between genetic variants of the major bovine milk proteins | 28 |
| Figure 10 | Production of vapour phase ionised molecules from a liquid source | 47 |
| Figure 11 | Schematic diagram of an electrospray mass spectrometer | 47 |
| Figure 12 | Raw and transformed electrospray mass spectra | 50 |
| Figure 13 (a) | Alkaline-urea PAGE of whole milk samples from individual cows | 58 |
| Figure 13 (b) | Acid-urea PAGE of whole milk samples from individual cows | 58 |
| Figure 13 (c) | Non-denaturing 'native' PAGE of whole milk samples from individual cows | 59 |
| Figure 14 | Non-denaturing 'native' PAGE of salt precipitate prepared β -LG | 66 |
| Figure 15 | Elution profile of casein proteins from S-Sepharose Fast Flow ion-exchange column | 68 |
| Figure 16 | Alkaline-urea PAGE of peak fractions collected from ion-exchange chromatography of whole casein | 69 |
| Figure 17 | Capillary electrophoresis of (a) whey protein standards, and β -LG purified from acid whey of a β -LG AA phenotype (b), a β -LG BB phenotype (c), a β -LG AB phenotype, a β -LG CC phenotype, and a β -LG BC phenotype | 72 |
| Figure 18 | Analysis of trypsin purity | 75 |
| Figure 19 | Typical RP-HPLC tryptic peptide profiles for the casein proteins α_{S1} -CN B, α_{S2} -CN A, β -CN A ¹ and κ -CN A | 80 |
| Figure 20 | Overlay of three α_{S1} -CN tryptic peptide maps | 82 |
| Figure 21 | Tryptic hydrolysis peptide map of standard α_{S2} -CN A | 84 |
| Figure 22 | FAB-MS of peak eluting at 16.4 minutes in RP-HPLC chromatogram of α_{S2} -CN A tryptic digest. | 84 |

| | | |
|-----------|---|-----|
| Figure 23 | Overlays of three tryptic hydrolyses of (a) β -CN A ¹ , (b) β -CN A ² and β -CN B | 87 |
| Figure 24 | Overlays of tryptic hydrolyses of β -CN variants A ¹ , A ² and B | 88 |
| Figure 25 | Overlay of three κ -CN A tryptic peptide maps | 89 |
| Figure 26 | Raw ESI-MS spectrum and transformed spectra (inset) of a sample of purified α_{S1} -CN | 93 |
| Figure 27 | Average individual mass determinations for α_{S1} -CN B samples | 94 |
| Figure 28 | Average individual mass determinations for each α_{S2} -CN A sample | 95 |
| Figure 29 | Average mass determinations for each β -CN sample analysed | 96 |
| Figure 30 | Average individual mass determinations for each κ -CN sample analysed | 97 |
| Figure 31 | Individual mass determinations for each β -LG sample analysed | 101 |
| Figure 32 | Multiple molecular weight species detected in β -LG samples | 103 |
| Figure 33 | RP-HPLC chromatograms of (a) Sigma β -LG, β -LG purified from acid whey from a β -LG AA phenotype cow (b), a β -LG BB phenotype cow (c), and heat denatured and DTT reduced β -LG from an AA phenotype cow | 111 |
| Figure 34 | Transformed ESI-MS spectra of (a) Sigma purified β -LG, β -LG purified from acid whey from a β -LG AA phenotype cow (b), a β -LG BB phenotype cow (c), a β -LG CC phenotype cow (d), and heat-denatured, reduced β -LG from an AA phenotype cow | 112 |
| Figure 35 | Initial steps of the Maillard reaction | 113 |
| Figure 36 | Transformed ESI-MS spectra of β -LG in whey prepared by ultracentrifugation of the milk from a β -LG AA phenotype (a) and a β -LG BB phenotype cow (b) | 117 |
| Figure A | FAB-MS analysis of α_{S2} -CN A tryptic hydrolysis RP-HPLC peak eluting at 18.1 minutes | 164 |
| Figure B | FAB-MS analysis of α_{S2} -CN A tryptic hydrolysis RP-HPLC peak eluting at 18.4 minutes | 165 |
| Figure C | FAB-MS analysis of α_{S2} -CN A tryptic hydrolysis RP-HPLC peak eluting at 18.7 minutes | 165 |

TABLES

| | | |
|---------|---|-----|
| Table 1 | Tabulated results of phenotyping by PAGE methods | 59 |
| Table 2 | Expected tryptic peptides from α_{S1} -CN B and their calculated mass | 78 |
| Table 3 | Theoretical tryptic α_{S2} -CN A peptides and their calculated mass . . | 80 |
| Table 4 | Additional chymotryptic-like peptides generated from the tryptic hydrolysis of α_{S2} -CN A | 82 |
| Table 5 | Expected peptides from a tryptic hydrolysate of β -CN A ² | 83 |
| Table 6 | Theoretical tryptic peptides from hydrolysis of κ -CN A | 85 |
| Table 7 | Comparison of milk protein masses determined experimentally and calculated from the primary sequence | 96 |
| Table 8 | Number of codons specifying a particular amino acid | 120 |

ABBREVIATIONS AND TERMINOLOGY

| | |
|-------------------|---|
| α -La | Alpha lactalbumin |
| α_{s1} -CN | Alpha S1 casein |
| α_{s2} -CN | Alpha S2 casein |
| β -CN | Beta casein |
| β -LG | Beta-lactoglobulin |
| κ -CN | Kappa casein |
| 2-D | Two dimensional |
| BSA | Bovine serum albumin |
| CE | Capillary electrophoresis |
| CHO | Carbohydrate |
| Da | Dalton |
| DAD | Diode array detector |
| DNA | Deoxyribose nucleic acid |
| ESI-MS | Electrospray ionisation - mass spectrometry |
| FAB-MS | Fast-atom bombardment mass spectrometry |
| HPLC | High performance liquid chromatography |
| IEF | Isoelectric focusing |
| IEX | Ion exchange |
| PAGE | Polyacrylamide gel electrophoresis |
| PCR | Polymerase chain reaction |
| RFLP | Restriction fragment length polymorphism |
| RP | Reversed phase |
| UHT | Ultra-high temperature |

In the context of this work the following terms are defined as:

Genotype - the genetic constitution of an individual organism. The genetic make up of an animal which codes for the synthesis of an individual milk protein.

Phenotype - the characteristics of an expressed milk protein as a result of the interaction of its genotype with its environment. Post-translational modifications such as phosphorylation and glycosylation frequently alter the gene product (protein) prior to expression in milk.

Electrophoretic variant - a protein variant that is able to be resolved from other genetic variants by electrophoretic techniques.

Silent variant - a protein variant with (an) amino acid substitution(s) that has no net effect on the overall electrical charge of the protein - generally undetected by standard electrophoretic techniques. Silent variants may occur when a neutral amino acid residue is substituted by another neutral residue, or when a residue carrying a charge is substituted by a similarly charged residue.