Journal of Cellular Biochemistry 49:121–127 (1992)

Prospects for the Genetic Engineering of Milk

A.J. Clark

AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Roslin, Midlothian EH25 9PS, United Kingdom

Abstract Milk and milk products comprise a substantial fraction of the protein intake of the industrialised West. The establishment of germline manipulation techniques in cows offers opportunities for directly manipulating milk composition to produce products with enhanced nutritional and processing properties. The major milk proteins are encoded by a small number of abundantly expressed single-copy genes and a number of possible manipulations are described. Milk proteins exhibit complex interactions with each other and with other constituents of milk. It will, therefore, be necessary to utilise model systems to evaluate the consequences of these proposed changes before embarking upon the costly and time-consuming process of manipulating the bovine genome. © 1992 Wiley-Liss, Inc.

Key words: milk/protein, transgenic, bovine, dairy industry

Milk is the sole food of the newborn mammal, and milk from domesticated livestock, such as cows, sheep, and goats, has been an important source of protein in the human diet since prehistoric days. Today it is estimated that about 30% of dietary protein in the industrialised West is derived from cow's milk [1]. In recent years the conventional genetic selection of dairy cattle has improved milk production dramatically. Additionally, some differences in milk composition between different breeds have been obtained. Germline manipulation of dairy animals offers the opportunity for the direct manipulation of milk composition and, indeed, this approach is being effectively exploited as a route for the production of biomedical proteins [2,3]. This review, however, will concentrate on the prospects for genetically engineering milk proteins for use in the dairy and associated food industries.

MILK

In the cow more than 90% of the total protein of milk comprises just six tissue-specific proteins which are synthesised by and secreted from the secretory epithelia of the mammary gland during lactation. These are the four caseins (α_{s1}, α_{s2}, β, and κ) and the two major whey proteins, β-lactoglobulin and α-lactalbumin. Whey also contains a variety of other proteins, including lactoferrin, immunoglobulins, and serum albumin [4]. The genes encoding the major milk proteins are single copy and they are expressed at high levels in a tissue-specific manner. They have been cloned and many of them are now completely characterised at the DNA sequence level. In principle, their genetic manipulation should be relatively straightforward. In practice, however, the situation is complicated by the fact that milk is not a simple solution of proteins. Thus, the major protein component, the caseins, are assembled into supra-macromolecular structures, termed micelles, which are in a colloidal suspension in milk (Fig. 1). It is the structure of the micelle that governs many of the complex properties of milk and its industrial uses. These properties are in turn governed by the properties of the individual caseins, as well as their interaction with each other and with other constituents of milk, such as calcium and whey proteins. Although there has been considerable research studying the biophysics and biochemistry of milk and milk proteins, it is an incompletely understood system and, at present, it is not possible to predict with certainty the precise consequences of many of the genetic modifications envisaged below.

MANIPULATION OF THE BOVINE GENOME

Most dairy products are derived from cows and, therefore, any proposal to manipulate milk composition pre-supposes that the genetic engineering of the bovine genome is a realistic prop-

Received March 3, 1992; accepted March 5, 1992.

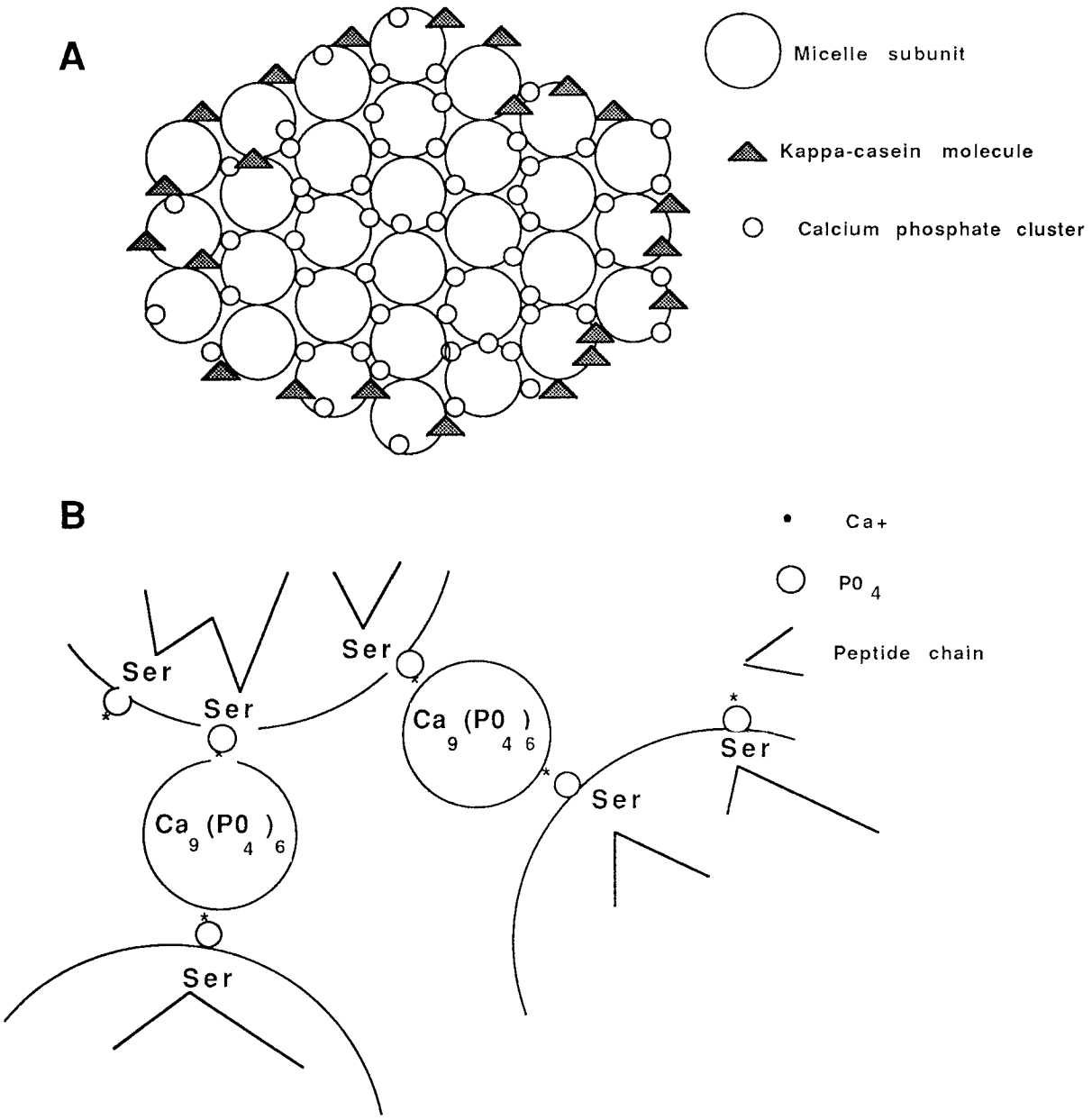

Fig. 1. Schematic representations of a casein micelle (**A**) and submicelles (**B**). Caseins exist in milk in roughly spherical structures, 100–300 nm in diameter, in the form of colloidal particle termed micelles. Approximately 10% of the micelles consists of calcium and phosphate with lesser amounts of citrate and magnesium. Micelles are comprised of submicelles that mainly contain α and β caseins. The micelle is stabilised by the binding of submicelles via calcium phosphate clusters and by the presence of κ casein molecules which are primarily localised at the surface [adapted from reference 5].

osition. Gene transfer by pronuclear injection has been established in a number of species of domestic livestock [6]. In pigs and sheep the success of producing transgenic animals is between 0.5% and 1% in terms of the number of eggs injected and transferred per live transgenic animal born. Nevertheless, the procedure is expensive, time consuming, and confounded by such factors as the small number, fragility, and opacity of pronuclear eggs and (in sheep) the limitation to litter size, which means that large numbers of recipient animals have to be employed. These factors are also encountered with the production of transgenic cattle, where a generation time of at least two years and the significantly greater carcase and husbandry costs

serve to magnify the problem. Some success has been attained and at least two live transgenic cattle reported [7,8]. In the latter study a major limitation to the generation of transgenic livestock was overcome by the use of in vitro cultured embryos derived from slaughterhouse ovaries. Other advances in embryo technology, such as screening cultured embryos for transgene insertion prior to transfer and the retrieval of large number of immature oocytes from live animals prior to in vitro fertilisation and culture, can also be expected to improve this technology.

In mice it is now possible to target specific mutations to endogenous genes using homologous recombination in embryo-derived stem (ES) cells [9]. An ability to target specific changes to endogenous milk protein genes (whether gene deletion or gene replacement) would provide a powerful approach for manipulating milk proteins. Despite intensive efforts and some encouraging results [10], ES cells which can colonise the embryo and contribute to the germ line have yet to be described for any species of livestock. If bovine ES cells are produced, they will revolutionise germline manipulation in this species. As well as enabling gene targeting, it seems likely that their use would supercede traditional pronuclear injection as a means of introducing new genes into the germline, particularly if nuclear transfer could be utilised to avoid the chimaeric generation [11].

Model Systems

Although gene transfer into cattle has been established, and the prospects for improving the technology are generally good, the time-scale and cost of these procedures, plus the complex biochemistry of milk proteins, means that model systems will be needed to initially evaluate many of the proposed changes.

Expression of milk proteins in prokaryotes such as *E. coli* or simple eukaryotes such as yeast has been attempted [12]. These expression systems are not capable of carrying out the appropriate post-translational modifications, such as phosphorylation and glycosylation, that determine many of the properties of milk proteins, so their general utility for evaluating the structure and function of genetically engineered derivatives is limited.

Established mammalian cell lines are widely used for the production of recombinant proteins and many efficient expression vectors are now available. There is, however, little published information on their use to produce and evaluate recombinant milk proteins. Preliminary studies indicate that it may be difficult to produce large amounts of the appropriately modified and secreted milk protein by this route [13,14]. Secretion of caseins from immortalised bovine mammary epithelial cells has been demonstrated [15]. These cells secrete only low amounts of milk proteins and require complex culture conditions. Nor do cultured cells produce milk, so it will not be possible to use them to study the behaviour of genetically engineered derivatives in their natural milieu, in particular their interaction with other milk proteins and (for caseins) their organisation into micelles.

Transgenic animal models provide a solution to some of these difficulties. A variety of milk protein genes have been introduced mice, including rat β-casein [16], bovine α-lactalbumin [17], ovine β-lactoglobulin [18], and rat whey acidic proteins [19], and expression obtained in the mammary gland. In some of these examples the foreign protein was secreted into milk at high levels. For β-lactoglobulin very high levels have been obtained and some lines of mice produce more than 20 mg/ml of the foreign protein in their milk (Fig. 2). The presence of large amounts of this foreign milk protein does not appear to affect the physiology of the mammary gland in terms of fat or carbohydrate metabolism, nor is the growth of the pups suckled on this milk compromised [20]. Interestingly, however, the total protein content of the milk from these animals is not increased, indicating that in mice, at least, there is a limit to the level of protein production from the gland. Although high-level β-lactoglobulin expression does not appear to compromise mammary function in-vivo, this may not be the case for other milk proteins. Thus, expression of the mouse WAP gene in transgenic pigs can inhibit normal mammary development and the gland appears unable to sustain normal lactation [21].

The transgenic mouse model is presently limited by the fact that the properties of mouse milk are poorly characterised. Since the ultimate aim is to manipulate cow's milk, this will involve extrapolation from a heterologous system and caution will have to be exercised. Given that milk is, broadly speaking, qualitatively similar between different species, mice should be useful for a preliminary evaluation of many of the changes. Nevertheless, the relatively small

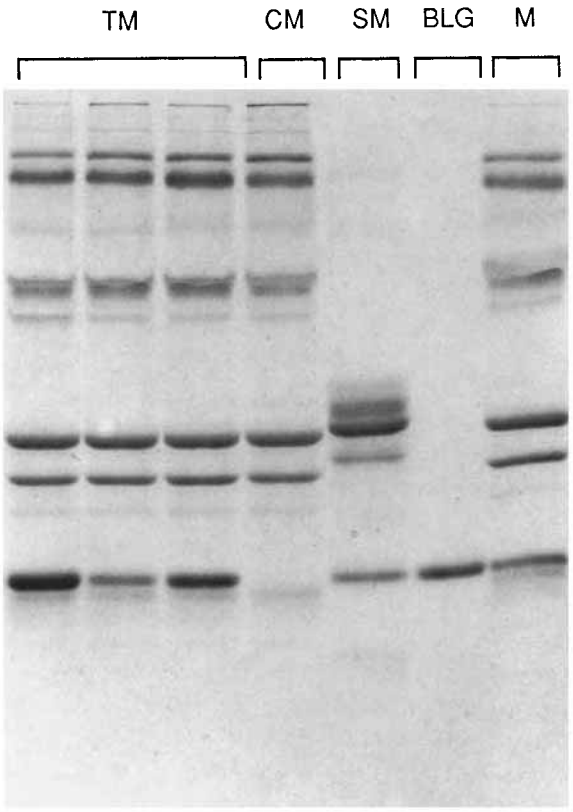

Fig. 2. SDS PAGE analysis of modified mouse milk. Milk samples from transgenic mice expressing the ovine beta-lactoglobulin gene (TM) were compared with sheep milk (SM) and control mouse milk (CM). BLG, purified sheep beta-lactoglobulin; M, markers. In some of the transgenic milk samples, the sheep beta-lactoglobulin is the most abundant protein. This high level expression has no apparent deleterious effects on the expressing female or upon the pups suckling the milk.

amounts of milk available (at most a few mls per lactation) may limit some of these analyses, particularly since many techniques of dairy chemistry have been established using substantial amounts of working material. In this regards other species, such as rabbits or sheep, may prove to be of more utility when it finally comes to modelling the behaviour of genetically engineered milk proteins (e.g., in cheese-making). Sheep, in particular, have the advantage that the overall composition of milk and the primary structure of the milk proteins are very similar to the cow (e.g., ovine and bovine β-lactoglobulin are both 162 amino acids long and differ at only five amino acid residues).

The Genetic Manipulation of Caseins

The caseins determine many of the properties of milk and upon them depend many of the industrial uses of milk and the milk proteins. The presence of additional α and β casein in milk enhances curd firmness and processing properties during cheese-making. Jimenez-Flores and Richardson [22] have estimated that a 20%–30% change in the level of a particular casein will profoundly affect these properties. This could be achieved simply by introducing additional casein genes into the bovine genome. These properties of casein are due to their phosphorylated status and so increasing the degree of phosphorylation may also have important consequences for the overall composition and processing properties of milk. In particular, since the phosphate groups bind calcium to form the micellar aggregates, increasing the phosphate content can be expected to enhance the overall calcium content, providing both a nutritional advantage as well as increasing the thermal stability of the micellar aggregates. Casein kinase, which is present in the Golgi apparatus, generally phosphorylates serines or threonines when these residues are specified by the amino acid motif Ser/Thr-X-Ac, where X can be any amino acid and Ac is an acidic amino acid residue, in particular Glu, Asp, or Ser-P. Additional serines with the potential for phosphorylation could be introduced into accessible regions of the molecule by introducing an extra coding segment. Alternatively, it should also be possible to recruit unphosphorylated serines by engineering the appropriate motif. For example, an additional phosphate centre could be created in β-casein at the sequence Ser_{164}-Leu-Ser-Gln-Ser-Lys-Val_{170} by changing Val_{170} to a Glu_{170}.

β-casein is 209 amino acids in length. It is rich in proline and contains 5 phosphorylated serines clustered in the first 35 amino acids at the N-terminal of the molecule. The cDNA sequence has been determined and the genomic structure elucidated. The gene is 8.7 kb in length and comprises 9 exons [23]. It is hydrophobic, but possesses a hydrophilic region, due principally to the clustering of acidic phosphoserine residues at the N-terminal. Because it is amphiphilic, it is an excellent emulsifier and this protein is used in products in which emulsions and food stability are required, such as cream liqueurs and coffee whiteners. One potential manipulation of β-casein is to alter the balance between the hydrophobic and hydrophilic portions of the molecule by changing the degree of phosporylation through the introduction of additional phos-

phoserines in the N-terminal region of the molecule.

α_{S1} and α_{S2} caseins are also extensively phosphorylated, and it will be possible to increase the phosphate content of these proteins by genetically engineering additional phosphoserine residues. α_{S1} casein is 199 amino acids in length and comprises 8 phosphorylated serines, 7 of which are clustered between residues 41 to 77 of the mature peptide. α_{S2} casein is 207 residues in size and, in contrast to α_{S1} and β casein, exists in a number of forms differing in their degree of phosphorylation (between 10 and 13 phosphate residues per molecule). The cDNA sequence of both α_{S1} [24] and α_{s2} [25] casein have been published. The corresponding gene structures have not yet been fully elucidated, although it is clear that they are more complex than the corresponding β-casein gene (A.G. McKinlay, personal communication) [25].

The proteolytic breakdown of the casein micelle occurs during the maturation of cheese, and the cleavage of a Phe_{24}-Val_{25} or Phe_{24}-Phe_{25} bond in α_{S1} casein is thought to be one of the key processes during ripening. Additional sensitive bonds would be expected to accelerate this process and Kang and Richardson [26] have proposed that the conversion of Ile_{71} to Phe_{71} in α_{S1} casein would generate an additional sensitive bond that would increase the rate of textural development during the ripening process.

κ-casein is 169 amino acids in length. It is phosphorylated to a lesser degree than the other caseins (it contains only one phosposerine per molecule) and, unlike them, it is O-glysoylated at a number of serine and threonine residues. κ-casein exists as a disulphide linked polymer comprising on average about 30 molecules in milk. The cDNA sequence has been published [24] and the corresponding gene comprises a 13 kb transcription unit divided into 5 exons [27]. By contrast to α and β caseins, κ casein is found predominantly on the surface of the micelle, where it determines micelle size and functions to prevent micellar aggregation. It is known that κ-casein plays an important role in preventing the thermally induced gelation of milk [28]. Expressing additional unmodified κ-casein genes will decrease micelle size and increase the stability of casein aggregates, retarding detrimental effects such as coagulation and gelation in various milk products.

In their simplest form these proposed changes to milk involve manipulating individual caseins by the introduction of the corresponding genes or their genetically engineered derivatives into the germline. As such, the consequences of these manipulations will have to be assessed in the presence of the endogenous caseins. Experiments in mice suggest that this type of manipulation will not result in any net increase in the protein content of milk, but, rather, will cause an overall qualitative change, such as increasing the proportion of casein to whey. For the future, if bovine ES cells are developed this will open up the possibility of manipulating endogenous genes specifically. It will be possible to target precise genetic changes and assess the properties of genetically engineered variants in the absence of their endogenous counterparts. The four casein genes are tightly linked in the bovine genome and contained within a segment of chromososmal DNA that is about 200 kbp in length [29]. It can be anticipated that gene targeting techniques will ultimately enable the specific deletion of large segments of chromosomal DNA. Thus it may be possible to delete the entire locus and replace it with a novel set of casein or indeed other genes for expression in milk! Finally, in mice, the major casein genes have also been shown to be tightly linked [30]. Deletion of the entire casein locus in this species would generate a useful model system for the subsequent testing of various genetically engineered caseins and combinations thereof.

The Genetic Manipulation of Whey Proteins

In the cow, whey proteins comprise about 18% of the total milk proteins. Although they are not generally considered to be of such relevance to the dairy industry as the caseins, they, nevertheless, exhibit important characteristics with regard to the dietary and functional properties of milk.

β-lactoglobulin is the major whey protein in cow's milk. It is a 162 amino acid globular protein. It has been proposed that β-lactoglobulin is involved in the transport of vitamin A, although the importance is this process is generally unclear, since milks from a number of species (e.g., human and rodent) lack this protein. The cDNA sequence has been published [31] and the corresponding genomic structure partially characterised [32].

The presence of β-lactoglobulin in milk causes a number of problems, both in terms of food processing and nutrition. The protein contains a free sulphydryl group at Cys_{121}. At higher tem-

peratures this group is exposed and is available to form disulphide bonds with free sulphydryl groups present on κ-casein, causing an unwanted gelation of milk. Secondly, the compact globular structure of β-lactoglobulin makes it difficult to digest. This is particularly a problem in infant formula milks in which the ratio of *whey to casein* has been increased and β-lactoglobulin is thought to be responsible for many of the observed allergies to cow's milk [33]. There is thus a case for the removal of β-lactoglobulin from milk and, should bovine ES cells become available, this would be practicable by gene targeting. Some care, however, must be exercised with this approach, since β-lactoglobulin is the main source of cysteine in milk. Therefore, on nutritional grounds the case is rather to replace β-lactoglobulin with a more digestible, less allergenic protein containing suitable amounts of cysteine.

α-lactalbumin is the other major whey protein and is 123 amino acids in length. The cDNA has been cloned and sequenced, and the genomic structure elucidated [34]. The transcription unit is 3 kb in length and comprises 4 exons. α-lactalbumin is one of the two subunits of the enzyme lactose synthase, which in the mammary gland is involved in the last step of lactose biosynthesis. Lactose is a disaccharide, comprising a glucose and a galactose moiety, and is present in nearly all mammalian milks. After ingestion, it is hydrolysed in the gut by lactase. In Indo-European populations the majority of people produce lactase throughout their life, but in African, Asian, and Indo-American populations, a post-weaning switch reduces its production. This results in lactose malabsorption, with consequent intestinal problems, and this excludes milk and many milk products from the diet of these people. Lactose malabsorpbers may represent up to 90% of the world's adult population [35]. A second problem with lactose is that its relatively high concentration in milk presents problems to various industrial processes in the dairy industry. The complete removal of lactose from milk by targeting the deletion of the α-lactalbumin gene (given that bovine ES cells will eventually be available) will not be desirable. This is because lactose is a major osmotically actively molecule in milk and functions to draw water across the secretory epithelium during lactation. Practically, then, the level of lactose can only be reduced rather than completely eliminated without compromising the functioning of the mammary gland. Anti-sense [36] and ribozyme approaches to reduce the steady-state level of α-lactalbumin mRNA have been suggested, but it is not clear how effective such approaches will prove. Alternatively, gene targeting to attenuate expression from the α-lactalbumin gene promoter or the introduction of a dominant/negative α-lactalbumin mutation could suppress lactase synthase activity and reduce the concentration of lactose.

Lactoferrin is an iron binding protein thought to play an important role in iron transport and absorption from the gut. This whey protein also has beneficial bacteriostatic properties. The concentration of lactoferrin is high in human milk, but relatively low in cow's milk. A receptor for lactoferrin is present on the mucosal surface of infant monkey small intestine [37]. Bovine lactoferrin, however, fails to bind to this receptor and this may explain the failure of the protein to enhance iron absorption in supplemented infant formula milks. Consequently, it may be necessary to express the human protein in milk and this approach is currently being developed. The human protein is 703 amino acids in length and the corresponding cDNA has been cloned and inserted into an expression cassette comprising regulatory sequences from the α_{s1} casein gene. The construct has been introduced into bovine germline by pronuclear injection and at least one transgenic animal produced [8]. At present no data is available on expression. This work represents the first attempt to alter bovine milk by gene transfer and has the advantage that a specific product and market (human lactoferrin supplemented infant formula milk) have been clearly identified.

CONCLUSION

The genetic engineering of cow's milk is a realistic proposition. Nevertheless, it will be necessary to evaluate and refine many of the changes proposed above using the appropriate model systems. Although some of the manipulations envisaged in this review do not yet have such a clearly defined target and market as, for example, human lactoferrin supplemented infant formula milk, this is not to say that such "designer milks," aimed at specific nutritional and/or industrial requirements, will not be increasingly exploited in the years to come. Nevertheless, the cost, in both time and money, of manipulating the bovine genome will always present limita-

tions to the adoption of this technology. Finally, the sensibilities of the consumers and the statutory bodies that exist to protect them will have to be carefully considered, given that these products are destined, ultimately, for human consumption.

REFERENCES

1. Hambreus L: In Fox PF (ed): "Developments in Dairy Chemistry—I." London and New York: Applied Science Publishers, 1982, pp 289–313.
2. Clark AJ, Bessos H, Bishop JO, Brown P, Harris S, Lathe R, McClenaghan M, Prowse C, Simons JP, Whitelaw CBA, Wilmut I: Bio/Techology 7:487–492, 1989.
3. Wright G, Carver A, Cottom D, Reeves D, Scott A, Simons JP, Wilmut I, Garner I, Colman A: Bio/Technology 9:830–834, 1991.
4. Jennes R, Sloan RE: Dairy Sci Abst 32:599–612, 1971.
5. Mepham TB: "Physiology of Lactation." Milton Keynes, U.K.: Open University Press, 1987.
6. Clark AJ, Archibald AL, McClenaghan M, Simons JP, Whitelaw CBA, Wilmut I: Proc NZ Soc Anim Prod 50:167–179, 1990.
7. Roschlau K, Rommel P, Anrewa L, Zackel M, Roschlau D, Zackel B, Schwerin M, Huhn R, Gazarjan KG: J Reprod Fertil (Suppl) 38:153–160, 1989.
8. Krimpenfort P, Rademakers A, Eyestone W, der Schan A, deBroek S, Kooiman, Kootwijk E, Platenburg G, Pieper F, Strijker R, de Boer H: Bio/Technology 9:844–847, 1991.
9. Capecchi, MR: Trends Genet 5:70–76, 1989.
10. Notarianni E, Laurie S, Moor RM, Evans MJ: J Reprod Fertil (Suppl) 41:51–56, 1990.
11. Smith L, Wilmut I: Biol Reprod 40:1027–1035.
12. Kang YC, Richardson T: J Dairy Sci 71:29–40, 1987.
13. Vilotte J-L: PhD thesis, University of Edinburgh, 1992.
14. Burdon, T: PhD thesis, University of Nottingham, 1988.
15. Huynh HT, Robitaille G, Turner, JD: Exp Cell Res 197:191–199, 1991.
16. Lee KF, Atiee SH, Rosen JM: Nucleic Acids Res 16:1027–1041, 1989.
17. Vilotte J-L, Soulier S, Stinnakre M-G, Massoud M, Mercier J-C: Eur J Biochem 186:43–48.
18. Simons JP, McClenaghan M, Clark AJ: Nature 328:530–532, 1987.
19. Bayna EM, Rosen JM: Nucleic Acids Res 18:2977–2985, 1990.
20. Wilde C, Clark AJ, Kerr M, Knight CH, McClenaghan M, Simons JP: Biochem J: (in press), 1992.
21. Wall RJ, Pursel VG, Shamay A, McKnight RA, Pittius CW, Hennighausen L: Proc Natl Acad Sci USA 88:1696–1700, 1991.
22. Jimenez-Flores R, Richardson T: J Dairy Sci 71:2640–2654, 1988.
23. Bonsing J, Ring JM, Stewart AF, MacKinlay AG: Aust J Biol Sci 41:527–537, 1988.
24. Stewart AF, Willis IM, McKinlay AG: Nucleic Acids Res 12:3895–3907, 1984.
25. Stewart AF, Bonsing J, Beattie CW, Shah F, Willis IM, McKinlay AG: Mol Biol Evol 4:231–241, 1987.
26. Kang YC, Richardson T: Food Technology 39:89–94, 1985.
27. Alexander LJ, Stewart AF, MacKinlay AG, Kapelinskaya TV, Tkach TM, Gorodetsky, SI: Eur J Biochem 178:395–401.
28. Fox PF: In Fox PF (ed): "Developments in Dairy Chemistry—I." London and New York: Applied Science Publishers, p 184, 1982.
29. Ferretti L, Leone P, Sgaramella V: Nucleic Acids Res 18:6829–6833, 1990.
30. Geissler EN, Cheng SV, Gusella JF, Housman, DE: Proc Natl Acad Sci USA 85:9635–9639, 1988.
31. Jamieson AC, Vandeyar MA, Kang YC, Kinsella JE, Batt CA: Gene 61:85–90, 1987.
32. Siva MC, Wong DWS, Batt CA: Nucleic Acids Res 18: 3051, 1990.
33. Saperstein, S. In Larson BL, Smith, VR (eds): "Lactation. A Comprehensive Treatise," Vol II. London: Academic Press, pp 250–280, 1974.
34. Vilotte J-L, Soulier S, Mercier J-C, Gaye P, Hue-Delahaie D, Furet, JP: Biochimie 69:609–620, 1987.
35. Dahlqvist A: In Delmont J (ed): "Milk Intolerance and Rejection." Basel: S. Karger, pp 11–16, 1983.
36. Mercier, J-C: In Smith C, King JWB, McKay, JC (eds): "Exploiting New Technologies in Animal Breeding." Oxford: Oxford Science Publications, pp 122–131, 1985.
37. Davidson LA, Lonnerdal B: Am J Physiol 254:G580–G585, 1988.