Journal of Cellular Biochemistry 49:333–338 (1992)

Isolation of Recombinant Proteins From Milk

Tracy D. Wilkins and William Velander

Transpharm/TechLab, Inc. (T.D.W.), and Department of Chemical Engineering, Virginia Polytechnic Institute (W.V.), Blacksburg, Virginia 24061

Abstract Milk is a complex bio-colloid which presents some unique problems for the protein isolation chemist, but the majority of the processing criteria for purifying recombinant proteins are the same as with any complex biological mixture. The casein micelles and fat globules behave as separate phases; they prevent filtration of the milk and interfere with the usual separation methods. The usual first step is to centrifuge the milk to remove the fat and precipitate the casein micelles with low pH or precipitating agents. Some recombinant proteins may associate to some degree with the micelles which may necessitate solubilizing them with chelating agents. If the majority of the product protein associates with either the fat or micelles, this can be used to advantage. Once the casein micelles have been removed or disrupted, the clarified milk can be processed by the usual separation methods. There also are proteases in milk which can degrade recombinant proteins. The greatest advantage of producing recombinant proteins in milk is the high concentration which can be obtained. The high levels of product protein can alleviate many problems associated with the application of classical purification strategies to transgenic milk proteins.

Key words: casein micelle, immunopurification, lactation, monoclonal antibody, post-translational modification, proteases, therapeutics, viral inactivation, whey

The requirements for purifying recombinant proteins are similar for source materials such as bacterial culture media, mammalian cells, or the milk of transgenic animals. A sequence of six major processing tasks is usually required: solids removal, volume reduction, pre-purification, viral inactivation, high resolution purification, and a final low resolution cleanup [see Velander et al., 1989; reviewed in Cartwright, 1987]. In addition, the protein chemist must be able to follow accurately the concentration of the protein through all the steps in the process as well as have reliable assays for the contaminants.

The recombinant protein must be in a reasonable concentration, proteolytic degradation must be controlled, and the process must produce the protein in a biologically active form. These requirements are not easy to accomplish in any system, but it is obviously easier in some than in others [reviewed in Marino, 1989]. With bacterial expression systems the recombinant protein may constitute 25% or even more of the total protein in the starting material; however, this protein may need to be solubilized, denatured, and refolded. The main problem with bacterial systems is that procaryotes cannot do the post-translational modifications which are frequently required for the biological activity of many eukaryotic proteins. Although recombinant proteins produced by eukaryotic expression systems may be made in a sufficiently native and soluble form, the recombinant protein is usually at very low concentrations [see Grinnell et al., 1990]. The use of milk as a vehicle for the expression of a recombinant protein has the advantages of high concentrations of soluble recombinant protein which has been post-translationally modified [see Wall et al., 1991; Wright et al., 1991; Denman et al., 1991]. However, the complexity of milk and the difficulty of isolating recombinant proteins from milk should not be underestimated.

MILK: A COMPLEX BIO-COLLOID

Milk is a complex bio-colloid and resembles blood plasma more than it resembles most other proteinaceous media [reviewed in Whitney, 1988]. The purification problems associated with milk [reviewed in Ladisch et al., 1989] are similar to that of plasma [see Velander et al., 1990] except that the levels of target protein will be much higher in milk. The level of background protein in milk is similar to that of blood plasma but the major constituents are very different.

Received March 18, 1992; accepted March 25, 1992.

The production of recombinant proteins in milk does not appear to reduce the concentration of the normal milk proteins. Transgenic milks containing recombinant protein at levels ranging from 1–65 g/l have been shown to be similar in levels of endogenous protein to that of nontransgenic milk [see Wall et al., 1991; Wright et al., 1991; Denman et al., 1991].

Milk is designed to give the newborn a source of fat, easily digested protein, and calcium [reviewed in Jenness and Sloan, 1970; Klobasa et al., 1987]. A family of small proteins, the caseins, bind calcium and keep it from precipitating; these proteins also are made to be easily cleaved by proteases to facilitate digestion by the neonate [Andrews and Alichanidis, 1983]. Caseins constitute over 50% of the protein in various milks, but there is some variation in the proportion of total caseins in various milks (Table I).

The caseins form aggregates with the calcium such that a colloidal suspension of protein and calcium results. These aggregates behave as micelles. The casein micelles are not filterable and can cause fouling in most devices [see Ladisch et al., 1989]. In addition, casein micelles cannot be separated from milk by industrial scale centrifugation and require ultracentrifugation. Because of this, the casein micelles usually are precipitated before centrifugation.

"Know thy enemy" is a must for the protein isolation chemist; the contaminants that are present in any starting material should be well characterized so that logical isolation procedures can be devised and so that the absence of these materials can be demonstrated in the final product. The electrophoretic protein profiles for various skim milks are shown in Figure 1. The major "whey proteins" are lactoglobulin and lactalbumin. Lactalbumin modifies galactosyltransferase to shift enzymatic activity from the transfer of galactose residues to glycoproteins to the production of lactose. Like caseins,

TABLE I. Milk Composition of Various Livestock*

Livestock	Total solids	Fat	Casein	Whey	Lactose
Cow	12.7	3.7	2.8	0.6	4.8
Goat	13.2	4.5	2.5	0.4	4.1
Sheep	19.3	7.4	4.6	0.9	4.8
Pig	18.8	6.8	2.8	2.0	5.5

*All values compiled from Jenness and Sloan [1970] and are given in weight percent.

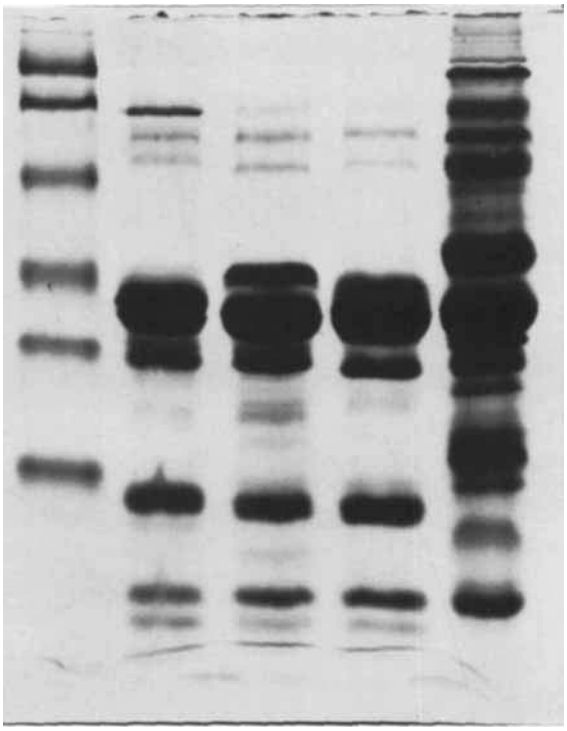

Fig. 1. Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) of skim milks from various livestock. SDS-PAGE is stained with Coomassie Blue dye. **Lane 1:** Molecular weight markers (106 Kda, 80 Kda, 49.5 Kda, 32.5 Kda, 27.5 Kda, 18.5 Kda). **Lane 2:** Skim milk from cow. **Lane 3:** Skim milk from goat. **Lane 4:** Skim milk from sheep. **Lane 5:** Skim milk from pig.

most milk proteins are small—lactoglobulin (M_r = 18,000) and lactalbumin (M_r = 14,000)—but the lactoglobulins in some species aggregate. These proteins exist in several different genetic forms and with different amounts and types of glycosylation patterns [see Stewart et al., 1987; Bonsing and Mackinlay, 1987]. The diversity and genetic variation of milk proteins can complicate purification procedures.

In addition to proteins synthesized by the mammary gland, plasma proteins can pass over into milk [see Klobasa et al., 1987; Whitney, 1988; Honkanen-Buzalski and Sandholm, 1981]. The concentration of serum passover proteins will vary during the lactation cycle and increase greatly if the mammary tissue becomes inflamed. Serum albumin (M_r = 58,000) is a major constituent of milk from healthy mammary glands and thus appears to be actively transported into milk from plasma. The immunoglobulins (M_r = 150,000–400,000) also are thought to originate in the plasma and then are specifically transported into milk. Immunoglobulins

occur at very high concentration in the early days of lactation (colostrum). There are many more plasma proteins in milk than can be visualized by staining with Coomassie Blue (Fig. 1). Indeed, most plasma proteins are found in milk in minor amounts. In addition, proteins released by the lysis of white blood cells and bacteria can be detected in most milks. Non-protein contaminants such as lipids, carbohydrates, and nucleic acids also are present in milk and must be removed during the purification of a recombinant product.

PROTEASES IN MILK

Milk maintains a delicate balance of protease activators and inhibitors and even inhibitors of activators [see Korycka-Dahl et al., 1983]. Therefore, proteolysis can occur spontaneously. Plasminogen, a serum protein, occurs naturally in milk where it can be converted into plasmin, which is a serine protease with very broad specificity. Plasmin is bound predominantly to the casein micelles which are preferred substrates for many proteases [see Schaar, 1985]. Thus, the high concentration of caseins may act to competitively inhibit proteolysis of recombinant proteins in the milk. There has not been enough experience with expression of different types of proteins in milk to determine whether milk proteases will be a significant problem. However, results from milk doping experiments using human plasma-derived Factor IX have suggested that protease activity is not a problem in the case of sheep milk [see Clark et al., 1989]. If protease activity is found to be a problem, inhibitors can be added at the same time that the milk is collected, but this would not protect the recombinant protein during the time it resides in the udder or mammary tissue. Frequent milking may be required to minimize this potential problem.

If proteolytic degradation of the transgenic protein occurs in vivo, it may be possible to co-express a protease inhibitor protein. In a few years, genes may be introduced only into special breeds of farm animals which express protease inhibitors and other protective molecules in their milk; the use of "wild type" farm animals may seem naive then, just as the use of wild type *E. coli* would seem today.

INTERFERENCE WITH ASSAYS

Protein assays on milk and milk fractions can give deceiving results. For example, caseins react differently than globular proteins with colorimetric protein assays [see Owen and Andrews, 1983]. The aggregation of the caseins into micelles also gives lower values than solubilized micelles, etc. The type of protein used as a standard also will influence this result (i.e., non-micellular casein or albumin). Proteose peptones (peptides) resulting from proteolysis of caseins are present in milk in varying amounts [Andrews and Alichanidis, 1983]. These peptides can interfere with rapid protein assays to give erroneously high results.

In the case of a recombinant enzyme, it is likely that one or more of the (over) 50 enzyme activities known to occur in milk will interfere with enzymatic assays of the transgenic protein. The level of these endogenous enzymes is not constant since many of them are the result of lysis of white blood cells or spontaneous activation [see Korycka-Dahl et al., 1983; Honkanen-Buzalski and Sandholm, 1981]. Thus, both the precision and accuracy of activity assays for the recombinant protein tend to be poor when performed on crude milk or milk derivatives.

Because of the limited effectiveness of most activity assays, immunological detection is often the only useable method. ELISA assays, Rocket Immunoelectrophoresis, and other quantitative immunological procedures can be used to follow purification. Care should be taken in the selection of antibodies used in these assays; the antibodies may recognize only a portion of the entire population of recombinant protein molecules. Several variants of the recombinant protein likely will be produced in milk, and some of these variants may be different from the spectrum of native molecules to which the antibody was made. Differences in glycosylation patterns and other processing can affect even the binding of monoclonal antibodies and also can impact the avidity of polyclonal antibodies. For example, a monoclonal ELISA that works perfectly for a human protein spiked into milk may only detect a portion of the recombinant protein molecules in milk. As a result, immunological assays using polyclonal antibodies may give better estimates of the levels of recombinant protein.

DIFFERENCES IN RECOMBINANT MOLECULES

The method of purification must address the likelihood that many slightly different forms of the recombinant protein will be present. Ideally, only the forms which are biologically active would be isolated. These different forms of the protein

may partition very differently during the isolation process and confuse the isolation chemist unless assays are available to differentiate the useful from the useless forms of the protein.

The main purpose of expressing a protein in milk is to obtain large amounts of protein with the modifications which can only be accomplished by mammalian cells [reviewed in Marino, 1989]. The diversity of product molecules can be related to the rate of protein synthesis. The rate of recombinant protein synthesis can exceed the rate at which the nascent peptide can be modified by mammalian cells. In the case where the protein synthesis rate exceeds the rate of post-translational modification, a significant portion of the recombinant protein molecules could be biologically inactive. The limits of glycosylation, proteolytic cleavage, gamma-carboxylation, etc., are not yet known for mammary tissue but have been documented in cell culture. For example, enzymatic saturation phenomena occurs in the gamma-carboxylation of glutamic acid residues by CHO cells in culture [reviewed in Grinnell et al., 1990].

There also will be differences in the number of sites glycosylated; even some normal milk proteins are heterogenous because of differences in the amount of glycosylation, as are normal plasma proteins. An additional problem is that in each species of dairy animal, different carbohydrates are preferentially attached to milk proteins. The incidence of attachment and nature of carbohydrate structure at each N- or O-linked glycosylation site will be highly dependent upon the secondary and tertiary structures that arise during post-translational processing [reviewed in Yan et al., 1989].

There has been insufficient experience with different types of proteins and levels of expression to know if over-loading of the mammary enzymes associated with post-translational modification will be a significant problem. Perhaps higher levels of the enzymes needed for post-translational modifications can be co-expressed via transgenics in the mammary cells.

METHODS FOR ISOLATION

The degree of difficulty in isolating the recombinant protein is directly related in most cases to the amount of the protein in the milk. In the case of alpha-1-antitrypsin, the extraordinary amount produced in sheep milk (30–65 g per liter) is about half the protein content of the milk, which makes the isolation task much less difficult [see Wright et al., 1991]. Functional forms of more complex proteins may be expressed at much lower levels and would be more difficult to isolate.

The colloidal nature of milk adds a degree of complexity not found with some other expression systems. Eliminating casein micelles is a high priority [see Ladisch et al., 1989]. The most common method is to lower the pH to 4.6; this causes immediate precipitation forming an acid curd and a whey which contains most soluble proteins. Unfortunately, low pH also can remove sialic acid residues from recombinant proteins. An alternative way to precipitate the casein micelles is with polyethylene glycol (PEG) at 15% or even with ethanol at moderate pH values. These precipitation methods are only useful if the recombinant protein partitions into either the precipitate or whey. For the case of a product which will distribute between phases, the casein micelles can be solubilized by chelation of calcium with EDTA or citrate [see Owen and Andrews, 1983]. The resultant solution can then be filtered, but the solubilized caseins must be removed at a later stage.

Once clarified, the whey can be subjected to classical chromatographic procedures. One of the most common methods is ion-exchange chromatography. The calcium and other ions can promote non-specific adsorption of proteins to the matrix, as well as alter the adsorption characteristics of the target protein. Diafiltration may be used to reduce the calcium concentration to overcome this problem.

Immunoaffinity chromatography is an attractive step for purification of high value proteins [see Velander et al., 1989]. Immunosorption is not easily applied to crude whey because of the high concentration of contaminating milk proteins. A large dilution of the milk would be required before application to the immunosorbent. Immunoaffinity could be used after removal of the majority of contaminating proteins by classical batch procedures. As mentioned above the monoclonal antibody used should bind all of the useful forms of the protein, and the elution must be gentle enough to retain the biological activity of the protein and minimize co-elution of contaminants and leaching of antibody. The latter problem necessitates a clean-up step to remove non-specifically bound contaminants and leached antibody molecules; ion-exchange chromatography is usually used for this step. Because of the high cost of immunosor-

bents, they must be regenerated for repeated use, but this can be a problem under the constraints of GMP (Good Manufacturing Practice).

PURITY

Therapeutic proteins may have different acceptable levels of "purity"; a protein administered only one time can be contaminated with small amounts of antigens which would not be acceptable in a protein which will be injected repeatedly over several years [reviewed in Ramabhadran, 1987]. Since there are no "pure" recombinant proteins, the assays used to detect contaminants are of major importance.

Gel electrophoresis (using conventional native and non-native conditions as well as isoelectric focusing) is a useful general guide to purity, but there can be many contaminating proteins present which will not be present at a high enough concentration for visualization, but will still be unacceptable in the final preparation [Ladisch et al., 1989]. Sensitive biological assays sometimes can detect such problem proteins, or immunological assays can be used. The protein also will have to be free of bacterial pyrogens (bacterial lipopolysaccharide) and free from detectable lipids, complex carbohydrate antigens, and nucleic acids [reviewed in Cartwright, 1987].

There must be a viral inactivation step in the purification scheme, and the animals must be maintained as a herd that is kept "virus free" with respect to those viruses that are known to be human pathogens. Lipid-envelope viruses can be effectively inactivated while other non-envelope and peptide pathogens are much more difficult to address [see Velander et al., 1990]. Rigorous quarantine of a production herd and their founder animals may also help minimize the incidence of human pathogens which have not yet been identified.

VARIATIONS BETWEEN MILKING

Isolation procedures are usually based on the premise that the starting material will be the same from day to day. In the case of transgenic milk, there is some variation to be expected in the milk and probably in the recombinant protein itself during the lactation cycle. The very first milk is the colostrum, which is very high in immunoglobulins and is very different from milk obtained in late lactation. During the bulk of lactation, the protein content is much lower than in colostrum, but can increase as milk volume decreases near the end of lactation. Inflammation of the teats, mastitis, is a major concern and should be closely monitored. Mastitis will cause large increases in the numbers of both bacteria and white blood cells in milk [see Honkanen-Buzalski and Sandholm, 1981]. Even subclinical infections could be of concern when recombinant proteins are being produced, so the milk should be checked at each milking for its acceptability for further processing.

Thus far, researchers have had insufficient experience with the expression of recombinant proteins in milk to know whether there also will be variations in the extent of post-translational modifications during the lactation cycle. Variation in the recombinant protein could occur even between animals with the identical gene construct. Since therapeutic proteins isolated from humans are also a mixture of types combined from thousands of people, variations in the population of product molecules may be an insignificant problem if the overall variation in the product is minimized by pooling milk from many different animals.

SUMMARY AND CONCLUSIONS

In this article we have tried to identify potential problems rather than emphasize the more obvious advantages of expressing a protein in milk. However, the recombinant protein will likely be at significantly higher concentration than that of the natural product. Thus, many potential contaminants may be effectively removed with classical processes which achieve a lower yield of product protein, but these processes may become economical with transgenic milk. In contrast, isolation processes for many natural products cannot sacrifice yield for purity due to the limited supply of starting material (i.e., human plasma). It is noteworthy that in spite of the many past and present difficulties which face the human plasma fractionation industry, it remains a viable industry for supplying a portion of many human therapeutic proteins. A similar chronology will undoubtedly be true for transgenic proteins isolated from milk.

REFERENCES

Andrews AT, Alichanidis E: J Dairy Res 50:275–290, 1983.
Bonsing J, Mackinlay AG: Dairy Research 54:447–461, 1987.
Cartwright T: Trends in Biotechnology 5:25–30, 1987.
Clark AJ, Bessos H, Bishop JO, Brown P, Harris S, Lathe R, McClenaghan M, Prowse C, Simons JP, Whitelaw CB, Wilmut I: Bio/Technology 7:487–492, 1989.

Denman J, Hayes M, O'Day C, Edmunds T, Bartlett C, Hirani S, Ebert KM, Gordon K, McPherson JM: Bio/Technology 9:839–843, 1991.

Grinnell BW, Walls JD, Gerlitz B, Berg DT, McClure DB, Ehrlich H, Bang NU, Yan, SB: In Bruley DF, Drohan WN (eds) "Advances in Applied Biotechnology Series, Protein C and Related Anticoagulants." Woodlands, TX: The Portfolio Publishing Company, 1990, pp 29–63.

Honkanen-Buzalski T, Sandholm M: J Dairy Res 48:213–223, 1981.

Jenness R, Sloan RE: Dairy Sci Abstr 32:599–612, 1970.

Klobasa F, Werhahn E, Butler JE: J Animal Sci 64:1458–1466, 1987.

Korycka-Dahl M, Ribadeau-Dumas B, Chene N, Martal J: J Dairy Sci 66:704–711, 1983.

Ladisch MR, Rudge SR, Ruettimann KW, Lin JK: In Fiechter A, Okada H, Tanner RD (eds) "Bioproducts and Bioprocesses." Berlin: Springer-Verlag, 1989, pp 209–221.

Marino MH: BioPharm 7–8:18–33, 1989.

Owen AJ, Andrews AT: Dairy Res 51:307–315, 1983.

Ramabhadran TV: Trends in Biotechnology 5:175–179, 1987.

Schaar J: J Dairy Res 52:369–378, 1985.

Stewart AF, Bonsing J, Beattie CW, Shah F, Willis IM, Mackinlay AG: Mol Biol Evol 4:231–241, 1987.

Velander WH, Madurawe RD, Orthner CL, Tharakan JP, Ralston AH, Strickland DK, Drohan WN: Biotechnology Progress 5:119–125, 1989.

Velander WH, Morcol T, Clark DB, Gee D, Drohan WN: In Bruley DF, Drohan WN (eds) "Advances in Applied Biotechnology Series Protein C and Related Anticoagulants." Woodlands, TX: The Portfolio Publishing Company, 1990, pp 11–27.

Wall RJ, Pursel VG, Shamay A, McKnight RA, Pittius CW, Hennighausen L: Proc Natl Acad Sci USA 88:1696–1700, 1991.

Whitney RM: In "Fundamentals of Dairy Chemistry." New York: Wong NP (ed) Van Nostrand Reinhold, 1988, pp 81–169.

Wright G, Carver A, Cottom D, Reeves D, Scott A, Simons P, Wilmut I, Garner I, Colman A: Bio/Technology 9:830–834, 1991.

Yan SB, Grinnell BW, Wold F: Trends in Biotechnological Sciences 14:264–268, 1989.