INVITED REVIEW

A Review and Proposed Nomenclature for Major Proteins of the Milk-Fat Globule Membrane 1,2

Ian H. Mather
Department of Animal and Avian Sciences, University of Maryland, College Park 20742

ABSTRACT

The characteristics and possible functions of the most abundant proteins associated with the bovine milk-fat globule membrane are reviewed. Under the auspices of the Milk Protein Nomenclature Committee of the ADSA, a revised nomenclature for the major membrane proteins is proposed and discussed in relation to earlier schemes. We recommend that proteins be assigned specific names as they are identified by molecular cloning and sequencing techniques. The practice of identifying proteins according to their Mr, electrophoretic mobility, or staining characteristics should be discontinued, except for uncharacterized proteins. The properties and amino acid sequences of the following proteins are discussed in detail: MUC1, xanthine dehydrogenase/oxidase, CD36, butyrophilin, adipophilin, periodic acid Schiff 6/7 (PAS 6/7), and fatty acid binding protein. In addition, a compilation of less abundant proteins associated with the bovine milk-fat globule membrane is presented.

(Key words: milk-fat globule membrane proteins, nomenclature, review)

Abbreviation key: ADPH = adipophilin, ADRP = adipose differentiation-related protein, BTN = butyrophilin, FABP = fatty acid binding protein, IgC1 = constant 1-type Ig domain, IgI = intermediate-type Ig domain, MDGI = mammary-derived growth inhibitor, MFGM = milk-fat globule membrane, MHC = major histocompatibility complex, PAS = periodic acid/Schiff, PAS 6/7 = periodic acid Schiff 6/7, VNTR = variable number tandem repeat, XDH/XO = xanthine dehydrogenase/oxidase.

The purpose of this report is to summarize the identities, characteristics, and sequences of the major proteins associated with the bovine milk-fat globule membrane (MFGM). This compilation follows a decade of progress in characterizing MFGM proteins through biochemical approaches, molecular cloning techniques, and computer-assisted sequence analysis. Since the last report from this Committee (104), complete amino acid sequences have become available for most of the major bovine proteins and for several MFGM proteins in other species.

We will discuss the sequences of bovine proteins in detail and use sequence comparisons with their respective homologs in other species to highlight conserved motifs and structural elements. An ancillary aim is to clarify misconceptions and problems of identity that have arisen in the literature because different groups have used different nomenclatures for the same proteins. Where appropriate, we will discuss the possible functions of MFGM proteins. However, it is not our intention to cover the origin and derivation of the MFGM or the regulation or mechanism of milk-fat secretion. Nor will we discuss the identity and protein composition of skim milk membranes (225, 380). These issues are covered elsewhere in a number of comprehensive reviews (11, 64, 205, 211, 215, 217, 254, 257, 259, 260, 267, 283, 307, 313, 314).

Isolation of MFGM

Bovine MFGM can be isolated by a simple four-step procedure in which the fat globules are separated from whole milk and washed several times with physiological buffers. The membrane is then released from the surface of the globules by physical or chemical means and collected by centrifugation [reviewed in (64, 104, 215, 217, 254, 267, 314)]. Bovine cream is most conveniently collected from large volumes of milk with a bench-top cream separator, whereas smaller volumes can be separated in standard laboratory centrifuges. Following separation, the cream fraction is typically washed two (258) or three (9, 37, 104, 231, 263, 267) times at room temperature with physiological buffers to remove skim milk proteins. The membrane is then released from the washed fat globules at reduced temperatures (<10°C) by physical agitation in a blender or homogenizer. Membrane also may be released by alternatively freezing and thawing the washed globules (216, 231, 250) or by direct extraction with nonionic detergents (308), bile salts (13, 152, 309), or polar aprotic solvents (94). The released membrane is then collected by high-speed cen-

Received July 22, 1999.
Accepted October 18, 1999.
e-mail: im2@umail.umd.edu.

1 Invited review under the auspices of the Milk Protein Nomenclature Committee of the American Dairy Science Association.

2 Because of the necessity of introducing gaps within sequence alignments, the numbered positions of the amino acids in the figures in this review do not match those in the referenced amino acid sequences. To get around this problem, the position of specific amino acid residues within the sequences are indicated in the text thus: -Xyz in which X is the single-letter abbreviation for the amino acid, y is the position number indicated in figures within this review, and z is the position number in the contiguous amino acid sequence (cited by reference).
trifugation, typically at 90,000 to over 100,000 × g for 1 h [e.g., (258)]. In some procedures membrane is precipitated at low pH [e.g., (181, 206)] or “salted out” with ammonium sulfate (160, 225) before centrifugation. All the major MFGM proteins are recovered in the precipitated membrane regardless of the method employed. However, there are considerable differences in the yields of specific components, especially the loosely adsorbed peripheral proteins, and the level of contamination from skim milk components (9, 11, 32, 37, 206, 215, 217, 263, 267).

To a large extent, the specific methods chosen for each stage are predicated by the needs of the investigator. However, some critical issues need careful attention. It is strongly recommended that the cream is separated from fresh, warm milk as soon as possible after milking (104) because fat globules and the MFGM are prone to structural changes as milk is cooled and during long-term storage (13, 26, 104, 195). For structural studies of the MFGM on intact fat globules, the method of Patton is recommended in which small amounts of cream are collected by flotation at low centrifugal force through a layer of buffer solution (311). The droplets suffer minimal damage during separation, and the topology of the major proteins is maintained [see, e.g., (28)]. For routine preparations, we recommend three washing steps with large volumes of buffer to remove skim milk protein (9, 13, 37, 104, 263, 267). Special attention should be paid to removing casein as completely as possible because the proteinase precursor, plasminogen, is tightly associated with casein micelles (37) and low levels of plasmin may be active in bovine milk (105). Casein and plasminogen can be reduced to very low levels in the washed fat globules by including physiological levels of salt in the washing buffers (32, 37). The inclusion of ε-amino caproic acid in the washing medium also reduces plasminogen levels in MFGM preparations by displacing plasminogen from casein micelles (37). In addition, proteolysis can be suppressed by the inclusion of proteinase inhibitors in the washing buffers. Elevated temperatures (37 to 40°C), sometimes used by investigators to release membrane entrained in the butter fat (258), should be avoided if possible.

Regardless of the method employed, the material released from the droplet surface (called here buttermilk) is separated by centrifugation into two major fractions: a soluble supernatant and a membrane pellet. From 5 to 20% of the globule-associated protein is recovered in the soluble fraction of bovine cream (206), depending on the method of isolation. Operationally, we will refer to the soluble fraction as the MFGM supernatant and to the membrane pellet as the MFGM.

Nomenclature

Historically, MFGM proteins were classified according to their relative mobilities during electrophoresis in polyacrylamide gels containing SDS. Bands of protein in the gels were identified by staining with Coomassie blue or the periodic acid/Schiff (PAS) reagent and assigned arabic or roman numerals. Typically, the protein bands were numbered consecutively from the top of the gel in order of increasing mobility and, therefore, decreasing Mr. Different numbering systems and different percentages of gels were employed by different investigators. These inconsistencies led to some confusion, especially with regard to the identity of the less abundant components and of different proteins with very similar mobilities. Problems also arose when MFGM proteins were compared between species, and mobility was used as the sole criterion for identity.

To circumvent some of these difficulties, the Milk Protein Nomenclature Committee recommended a classification scheme that specified the mobility, Mr, and staining characteristics of individual proteins and the percentage of polyacrylamide gel employed (32, 104). However, this proposal was never widely adopted. We recommend that this scheme and, where possible, the practice of identifying MFGM proteins by number be discontinued. As sequence data become available, specific proteins can be unambiguously identified and assigned appropriate names. A list of currently recommended names, abbreviations, and previously used names for the major bovine MFGM proteins is summarized in Table 1.

Major Proteins of Bovine MFGM and MFGM Supernatant Fractions

Bovine MFGM is resolved into 7 to 8 major bands of protein when separated by SDS-PAGE (Figure 1a, b, c, lanes 1, 3, 5). Most of these bands comprise single, major protein components (i.e., gene products), because they are each resolved into a series of related isoelectric variants when they are separated by two-dimensional gel electrophoresis. However, in nearly every case, minor unrelated proteins with the same apparent Mr, can be detected in the two-dimensional map when a sensitive silver stain is used (Figure 2).

The major proteins resolved are, in order from the top of one-dimensional SDS gels, the mucin MUC1 (310, 367), the redox enzyme xanthine dehydrogenase/oxidase (XDH/XO) (249, 263, 416), a poorly characterized glycoprotein PAS III (204), CD36 (134, 136), butyrophilin (BTN) (116, 189, 257), a complex series of bands comprising glycosylated variants of Periodic acid/Schiff 6/7 (PAS 6/7) (181, 382), and adipocyte differentiation-related protein (ADRP) (154) or adipophilin (ADPH)
Table 1. Recommended names and alternatives for the most abundant milk-fat globule membrane proteins.

<table>
<thead>
<tr>
<th>Recommended name (Abbreviation)</th>
<th>Alternative names&lt;sup&gt;1&lt;/sup&gt;</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mucin 1 (MUC1)</td>
<td>Glycoprotein A</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Nonpenetrating glycoprotein NPGP&lt;sup&gt;2&lt;/sup&gt;</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>NCRC1&lt;sup&gt;1&lt;/sup&gt;</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Polymorphic epithelial mucin PEM&lt;sup&gt;2&lt;/sup&gt;</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Ca antigen&lt;sup&gt;2&lt;/sup&gt;</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>MAM-6&lt;sup&gt;2&lt;/sup&gt;</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>Peanut-lectin binding urinary mucin PUM&lt;sup&gt;2&lt;/sup&gt;</td>
<td>209</td>
</tr>
<tr>
<td></td>
<td>H23&lt;sup&gt;2&lt;/sup&gt;</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>Glycoprotein 2</td>
<td>231, 367</td>
</tr>
<tr>
<td></td>
<td>DF3&lt;sup&gt;2&lt;/sup&gt;</td>
<td>234</td>
</tr>
<tr>
<td></td>
<td>Episialin&lt;sup&gt;2&lt;/sup&gt;</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td>Component I</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band I also Band II (?)</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>Epithelial membrane antigen EMA&lt;sup&gt;2&lt;/sup&gt;</td>
<td>298</td>
</tr>
<tr>
<td></td>
<td>PAS1</td>
<td>356</td>
</tr>
<tr>
<td></td>
<td>PAS O&lt;sup&gt;2&lt;/sup&gt;</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td>HM glycoprotein C&lt;sup&gt;2&lt;/sup&gt;</td>
<td>360</td>
</tr>
<tr>
<td>Xanthine dehydrogenase</td>
<td>Band I</td>
<td>9, 12</td>
</tr>
<tr>
<td>/oxidase (XDH/XO)</td>
<td>Component II</td>
<td>225, 231</td>
</tr>
<tr>
<td></td>
<td>Component 3</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band 2</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CB1</td>
<td>356</td>
</tr>
<tr>
<td>Periodic acid Schiff III (PAS III)</td>
<td>Glycoprotein C (?)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Glycoprotein 4 (?)</td>
<td>231</td>
</tr>
<tr>
<td></td>
<td>Component II</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>PAS3</td>
<td>356</td>
</tr>
<tr>
<td>Cluster of Differentiation (CD96)</td>
<td>Band III/Glycoprotein D(?)</td>
<td>9, 12</td>
</tr>
<tr>
<td></td>
<td>PAS IV</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>Component III (?)</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>Component III/IV/Glycoprotein 5 (?)</td>
<td>231</td>
</tr>
<tr>
<td></td>
<td>Component 10/III or 11 (?)</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band 3/IV</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CB4/PAS 4</td>
<td>356</td>
</tr>
<tr>
<td>Butyrophilin (BTN)</td>
<td>Band IV/Glycoprotein E</td>
<td>9, 12</td>
</tr>
<tr>
<td></td>
<td>Component IV</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>Component V/Glycoprotein 6</td>
<td>231</td>
</tr>
<tr>
<td></td>
<td>Component 12/IV</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band 4/V</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CB5/PAS 5</td>
<td>356</td>
</tr>
<tr>
<td>Adipophilin (ADPH)</td>
<td>ADPR&lt;sup&gt;2,3&lt;/sup&gt;</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>Component 14 (?)</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band 6 (?)</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CB6 (?)</td>
<td>356</td>
</tr>
<tr>
<td>Periodic acid Schiff 6/7 (PAS 6/7)</td>
<td>Band V/Glycoprotein F</td>
<td>181, 224</td>
</tr>
<tr>
<td></td>
<td>MGP 57/53</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Glycoprotein B</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Bovine-associated mucoprotein (BAMP)</td>
<td>72, 327</td>
</tr>
<tr>
<td></td>
<td>BA46&lt;sup&gt;2&lt;/sup&gt;</td>
<td>91, 237</td>
</tr>
<tr>
<td></td>
<td>p47&lt;sup&gt;4&lt;/sup&gt;</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>GP-55&lt;sup&gt;5&lt;/sup&gt;</td>
<td>200, 255</td>
</tr>
<tr>
<td></td>
<td>Component VI</td>
<td>225, 231</td>
</tr>
<tr>
<td></td>
<td>Components 15/16/V</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Band 5/VI and 7/VII</td>
<td>262</td>
</tr>
<tr>
<td></td>
<td>CB7 &amp; 8 (PAS 6 and 7)</td>
<td>356</td>
</tr>
<tr>
<td></td>
<td>MFG-E8&lt;sup&gt;3&lt;/sup&gt;</td>
<td>382</td>
</tr>
<tr>
<td></td>
<td>Lactadherin&lt;sup&gt;3&lt;/sup&gt;</td>
<td>394</td>
</tr>
<tr>
<td>Fatty-acid binding protein (FABP)</td>
<td>Mammary-derived growth inhibitor (MDGI)</td>
<td>46</td>
</tr>
</tbody>
</table>

<sup>1</sup>Name for bovine protein unless otherwise indicated.
<sup>2</sup>Human homolog.
<sup>3</sup>Mouse homolog.
<sup>4</sup>Pig homolog.
<sup>5</sup>Guinea-pig homolog.
<sup>?</sup>indicates some uncertainty because the assignment was made from photographs, scans, or drawings of gels.
Figure 1. Separation of bovine milk-fat globule membrane (MFGM) proteins by SDS-PAGE. Bovine MFGM (M) (lanes 1, 3, and 5) and MFGM supernatants (S) (lanes 2, 4, and 6) were separated in 8% (wt/vol) (a, b; lanes 1 to 4) or 12% (wt/vol) (c; lanes 5 and 6) polyacrylamide, as indicated in the figure. Gels were stained either with Coomassie blue (CB) (a, c; lanes 1, 2, 5, and 6) or the periodic acid/Schiff reagent (PAS) (b; lanes 3 and 4). Mr protein standards are indicated to the right of Panel a and were from top to bottom of the gel, respectively, rabbit skeletal muscle myosin (M, 200,000), Escherichia coli β-galactosidase (M, 116,250), rabbit muscle phosphorylase β (M, 97,400), BSA (M, 66,200), hen egg-white ovalbumin (M, 45,000), and bovine carbonic anhydrase (M, 31,000). The major proteins and glycoproteins discussed in the text are indicated to the left of Panel a and to the right of Panels b and c. The proteolytically clipped fragment of butyrophilin (BTN) lacking the C-terminus is indicated by a small arrow to the left of Panel a, lane 1. XDH/XO = xanthine dehydrogenase/oxidase, PAS 6/7 = periodic acid Schiff 6/7, ADPH = adipophilin, and FABP = fatty acid-binding protein.

(153) (Figure 1, a, b, c, lanes 1, 3, 5). A band of material associated with the dye front comprises several proteins and is of variable composition depending on the percentage of gel employed. Fatty acid binding protein (FABP) (46, 53) is a component of this material in gels containing ≤10% polyacrylamide (compare Figure 1a, lane 1, and Figure 1c, lane 5).

Of the major proteins, six stain strongly with Coomassie blue (XDH/XO, CD36, BTN, ADPH, PAS 6/7, and FABP), and they are the proteins most often illustrated by investigators in publications (Figure 1a, lane 1, and Figure 1c, lane 5). Two glycoproteins, MUC1 and PAS III, do not stain with Coomassie blue but can be detected with either the PAS reagent (Figure 1b, lane 3) or a modified silver-staining procedure (271) (see Figure 3 for example of MUC1). Many minor components are also detectable with the silver reagent. These will not be considered in this report.

The MFGM supernatant fraction is less well characterized, despite the fact that it may comprise over 20% of the total globule-associated protein. Identification of the component proteins is based largely on comparison of electrophoretic mobilities, staining characteristics,
and reaction with a limited number of specific antibodies [e.g., (53, 193, 206), I. H. Mather, unpublished observations]. Major components of this fraction include MUC1, XDH/XO, PAS 6/7, and FABP (Figure 1a, b, c, lanes 2, 4, 6). XDH/XO, PAS 6/7, and FABP lack membrane-spanning domains and are regarded as peripheral proteins of MFGM (181, 258, 260, 263), which are released into the aqueous phase as the membrane bilayer is stripped from fat globules during phase inversion. It is assumed that these soluble proteins are identical to the MFGM-associated forms, although this possibility has never been rigorously tested.

In the following sections we will review each of the major proteins identified in Figure 1 and conclude with a brief discussion of some of the less abundant proteins associated with MFGM.

**MUC 1**

The MFGM of most ruminant, rodent, and primate species analyzed are characterized by the presence of one or more heavily glycosylated mucin-like glycoproteins of high Mr (254, 310). Because of the large content of carbohydrate, these proteins typically stain well in gels with the PAS reagent (231, 258, 262, 359, 367) and modified silver stain (316) but poorly, if at all, with Coomassie blue. The predominant mucin in bovine MFGM is structurally related to the human mucin, MUC1 (310, 372). Hallmarks of MUC1 mucins across species are the presence of single-membrane anchors of Type 1 topology, extensive heavily glycosylated extracellular domains that contain a variable number of tandem repeats in the peptide sequence and short cytoplasmic tails (123, 310, 372).

Bovine MUC1 can be readily identified in the electrophoretic profile of MFGM proteins as PAS- or silver-stained bands of protein that migrate with slower mobilities than XDH/XO in SDS-polyacrylamide gels (Figure 1b, lane 3; Figure 3). Because of allelic polymorphism, MUC1 may appear as a diffuse band of stained material in pooled milk samples from several cows or as one or two discrete bands from individual animals (315, 316). In heterozygous animals, each allele encodes a 20-amino acid tandem repeat sequence of specific length [in humans, 41 and 85 repeats are most common in a Northern European population with extremes of 21 and 125 repeats (122)]. The alleles are expressed in a codominant manner, such that each gives rise to equal quantities of protein (144, 384). The two polymorphic forms in heterozygotes are readily resolvable by electrophoresis in low percentage SDS-polyacrylamide gels (4 to 6%, wt/vol) (312, 315, 316, 373) (Figure 3). MUC1 from individual homozygotes appears as a single band. Five alleles have been identified in US Holstein cattle, which give rise to MUC1 forms ranging in apparent Mr from approximately 160,000 to 200,000 in MFGM (159, 180). Of these, MUC1 forms of Mr 177,000 and 189,000 are the most common and predominate in the pooled milk of commercial herds (180). In a limited survey of other breeds, the MUC1 forms of Guernsey cows appeared most similar to those of Holsteins (315). Jersey, Ayrshire, and Brown Swiss cattle express two apparently different predominant forms with Mr of 170,000 and 200,000 (315). A total of five MUC1 alleles ranging in Mr from 120,000 to 220,000 have been identified in five Italian breeds (338).

MUC1 offers a convenient and accessible polymorphic marker for genotyping production animals at both the genomic and protein levels (159). Variable number tandem repeats (VNTR) are present throughout the genome, typically in noncoding regions. The VNTR in the MUC1 gene and other mucin genes (123) are rare examples of polymorphic sequences within exons that are transcribed and translated into proteins (384).

Bovine and human MUC1 have been referred to by many names that denote relative electrophoretic mobility or reactivity with monoclonal antibodies (Table 1). In retrospect, the use of electrophoretic mobility as a criterion was unfortunate because in some cases, polymorphic forms of MUC1 in MFGM samples from heterozygous animals were numbered differently on the assumption that they were different glycoproteins [e.g., Band I and Band II of reference (262) were probably different forms of MUC1]. Descriptive names have also been devised that emphasize the high content of sialylated glycans in MUC1 or specific expression in epithe-
Bovine MUC1 is associated with the cream fraction or skim milk membranes at an estimated combined level of up to 40 mg/L of milk (310); human milk contains substantially more [729 to 805 mg/L in the milk of 41 subjects (322)]. MUC1 is assumed to be membrane bound in freshly secreted milk, and analysis of caprine milk showed that it is associated with fat and skim milk membrane fractions (76). However, the integral membrane-bound form can behave as a soluble protein under certain conditions, and various soluble forms of MUC1 lacking the membrane anchor have been described.

MUC1 possibly behaves ambiguously as a soluble protein, because, in some species, it is associated with filamentous structures on the outer cell surface and secreted milk-fat globules (66, 310, 417). These filaments have been detected by freeze-etch electron microscopy on human and horse fat globules (66, 323, 417) and extend from 0.5 to 1 µm in length from the globule surface. Treatment of human fat globules by brief exposure to hot (80°C) water dislodges the filaments, and MUC1 can be recovered in the supernatant as a “soluble” protein (66). It is possible that a similar process may occur at a slower rate at physiological temperatures or during the processing of milk for analysis. For practical purposes, human MUC1 has been purified in soluble form from skim milk or acid whey by several investigators [e.g., (298, 360)]. Recent preliminary data suggest that a substantial amount of MUC1 in bovine milk that has been collected and stored overnight in a bulk tank is recovered in the aqueous milk serum phase (323). These data suggest that cooling or stirring are sufficient to release the mucin (323), possibly in filamentous form, from the surface of the MFGM or skim milk membranes [note that unpublished data of Buchheim cited by Peterson et al. (323) suggest that filaments are present on the surface of bovine fat globules, despite earlier reports to the contrary (66, 417)]. A similar mechanism may explain why such a surprisingly large amount of MUC1 remains in the MFGM supernatant fraction when bovine MFGM is recovered from buttermilk (compare lanes 3 and 4, Figure 1b).

MUC1 from several species behaves anomalously when fractionated in biphasic solutions of Triton X-114. Unlike most integral membrane proteins (50), MUC1 is largely recovered in the aqueous phase when Triton X-114 extracts of MFGM are induced to undergo phase inversion above the cloud point (134, 200). These properties are assumed to be intrinsic to the transmembrane integral form of MUC1. Presumably the heavily glycosylated, hydrophilic, exoplasmic domain masks the hydrophobic membrane anchor and facilitates partitioning into the aqueous layer.

Soluble forms of human MUC1 are secreted from breast tumor cells, and MUC1 has been detected in the sera of both normal individuals and cancer patients (70). MUC1 is also a normal component of human urine (209, 384). Some of these forms may be the transmembrane integral mucin, shed in membrane fragments or filaments (?) from epithelial cell surfaces. However, soluble forms of MUC1 lacking the membrane anchor have been described. These variants arise either from proteolytic cleavage at site(s) in the exoplasmic domain close to the membrane anchor (51, 163, 242, 310) (Figure 4b) or from splice variant mRNA (366). Experiments with cultured cell lines (162, 242) have shown that proteolytic cleavage within the exoplasmic domain generates a heterodimer in which the bulk of the exoplasmic domain is bound by noncovalent interactions to the membrane-associated C-terminal fragment (242, 310). A fraction of the exoplasmic domain may also be secreted into the medium (51, 163). The extent that this processing reaction occurs in vivo, in any species, to give rise to heterodimeric membrane-bound forms and secreted soluble variants is currently unclear. In at least one case, a soluble human form of MUC1 arises from a variant mRNA in which the open reading frame comprises the first two exons and the 5′ region of the adjoining intron (242). This secreted form has been detected in culture supernatants of breast cancer cells in vitro and in the sera of breast cancer patients (366).

Despite a wealth of data on physical and chemical characteristics, the biological functions of MUC1 are uncertain. MUC1 is expressed on the apical plasma membrane of epithelial cells in many tissues (16, 52, 202, 254, 306, 310) and is incorporated into the MFGM during the budding and release of lipid droplets at the cell surface. The exoplasmic domain physically dominates the apical surface and may extend from 0.2 to 0.5 µm above the membrane bilayer (163, 310, 418). As noted above, the MUC1-containing filaments are major external structures from 0.5 to 1 µm in length (66, 323, 417). In this capacity, MUC1 is presumed to protect exposed surfaces from physical damage and invasive pathogenic microorganisms (310, 323, 347). In addition, MUC1 may play immunoprotective roles in the suckling neonate by binding to and sequestering pathogenic mi-
Figure 4. Structure and sequence comparisons of MUC1. (a) Domain structure of MUC1, showing the signal sequence (SS) and tandem repeat (TR) in the exoplasmic domain (exo), the transmembrane anchor (TM), and short cytoplasmic tail (cyto). C = C-terminus; N = N-terminus. Diagram drawn to scale with 12 tandem repeats. (b) Comparison of the sequences of human, gibbon, mouse, rabbit, hamster, rat, guinea pig, and bovine MUC1 using the alignment program PILEUP [Wisconsin Package Version 9.1. Genetics Computer Group (GCG), Madison, WI]. Only partial sequences of rat, guinea pig, and bovine MUC1 are available, as shown in the figure. Alignments were made with a human MUC1 sequence containing 12 tandem repeats as indicated, and sequences were edited manually. The signal sequence and cleavage site (↓) and the proteolytic cleavage site and membrane anchor are indicated above the alignments. Key: Potential N-linked glycosylation sites in human and gibbon MUC1 (♦), conserved YYQEL sequence (*), conserved CQC sequence (†/H17006), glycogen synthase kinase 3β phosphorylation site (‡/H17034), and adjacent β-catenin–binding motif in human MUC1 discussed in the text (•). Identical residues are indicated by dashes below the human MUC1 sequence, and gaps in the alignment are indicated by dots. GenBank®/EMBL Accession numbers were as follows: human, J05581; gibbon, L41589; mouse, M65132; rabbit, U85787 and L41544; hamster, U36918 and L41545; rat, AF007554; guinea pig, L14546; and cow, L14543. (c) Comparison of the available consensus tandem repeat sequences from human, gibbon, mouse, rabbit, hamster, and cow.

croorganisms within the gut lumen (323, 347). It has also been suggested that MUC1 functions in the development of ducts and cavities during embryogenesis (52), possibly by reducing adhesive interactions between epithelial cell surfaces by steric hindrance (418). However, mice lacking a functional MUC1 gene develop and reproduce normally with no obvious alteration in phenotype (374). Human MUC1 has been the subject of intensive investigation because it is expressed at high levels in mammary carcinomas in both normal and aberrantly glycosylated forms, which may modulate the immune response toward tumor progression [for reviews see references (40, 123, 163, 173, 265, 396)].

**Sequence Analysis**

Complete cDNA and gene sequences are available for human (122, 243, 424), gibbon (372), and mouse (373) MUC1 and either complete or partial cDNA for MUC1 from cow, rabbit, hamster, guinea pig (161, 304, 372), and rat (GenBank Accession No. AF007554). Alternatively spliced variant mRNA encoding the human mucin have been described (243, 296, 424, 438).

The canonical structure of MUC1 comprises an N-terminal, cleavable signal sequence, a repetitive domain flanked on either side by nonrepetitive sequences toward the N-terminus and the membrane anchor, a single transmembrane domain, and a short cytoplasmic tail (Figure 4a). In primate MUC1 (human and gibbon), the regions flanking the tandem repeat are similar to the repetitive domain itself. These similarities decline further, away from the repetitive domain, until any resemblance to the repeat sequence is lost (122, 144, 243, 372, 424).

The polypeptide chain of human MUC1 has an estimated Mr of 49,224, including the N-terminal signal sequence and one 20-amino acid repeat sequence (122, 243, 424). Because the VNTR has an Mr of 1887, and it may be repeated from 21 to at least 125 times in individual molecules, the predicted Mr, of the human polypeptide may vary from approximately 87,000 to 283,000. With addition of large quantities of carbohydrate (50%, wt/wt), the Mr of fully processed MUC1 will likely range from 180,000 to over one-half million. In stein cattle (159, 180) suggest that MUC1 molecules in this breed have between 25 to 35 tandem repeats. These repeats will contribute approximately 100,000 to 145,000 Da to the total mass of individual molecules (Mr, range from 156,000 to 193,000 (180)).

Comparison of available sequences across species shows, as might be expected, that the human and gibbon mucins are most similar with an amino acid identity of approximately 93% (372, Figure 4b). Across mammalian orders, the most conserved regions are the transmembrane and cytoplasmic domains with amino acid identities around 90% (Figure 4b). An anti-peptide antibody was generated against the C-terminal 17 amino acids of human MUC1 that showed broad specificity for MUC1-like mucins across at least eight other species (319). Notable motifs within the cytoplasmic domain include the sequence, CQC (arrowheads, Figure 4b), contiguous with the transmembrane anchor and conserved in eight species, which may function in protein-protein interactions (437) and which appears to be necessary for expression of MUC1 on the cell surface (320).

The cytoplasmic domain also contains several potential phosphorylation sites on S (241), T, or Y residues (161, 304, 372, 437). The S674 residue (122, Figure 4b) is phosphorylated by glycogen synthase kinase 3β in human breast cancer cells (241). This modification may modulate associations between MUC1 and the cytoskeleton because increased phosphorylation of S674 inhibits binding of the cytoskeletal protein β-catenin to the adjacent sequence S680 to S689 (open circles, Figure 4b) (dots, Figure 4b). β-catenin functions as a linker protein between cadherin molecules in adherens junctions on the cell surface and actin microfilaments in the cytoplasm. Thus phosphorylation of a single S residue in the cytoplasmic tail may regulate associations between MUC1 and the actin cytoskeleton (241, 306).

Several Y residues within the cytoplasmic tail of MUC1 are within consensus sequences for interaction with the SH2 domain of signaling kinases and linker proteins, suggesting that MUC1 may function as a signaling receptor (372, 437). A spliced-variant of human MUC1 lacking the exoplasmic repetitive domain MUC1/Y (438) has been suggested to function as a cytokine-like receptor ([437]; however, see also reference (372)). Results of preliminary experiments show that Y residues in both MUC1 and MUC1/Y are phosphorylated in transfected cells in culture and that phosphory-
lated forms of MUC1 bind to SH2-domain-containing proteins in vitro (437).

Across orders, the most dissimilar region in MUC1 lies within the exoplasmic domain before the membrane anchor (Figure 4b). Strikingly, within this region, the sequence YYQEL (asterisks, Figure 4b) is absolutely conserved, at position, in seven species. This motif may function in membrane recycling or sorting of MUC1 to the apical plasma membrane (372).

The region of MUC1 of clearest functional significance is the repetitive domain, which dominates the exoplasmic sequence. Although there is considerable variation among the tandem repeat sequences, several key features have been conserved across species. The average length of the repeat sequence is 20 amino acids, and it is rich in S, T, P, and A residues. One S and two P residues have been absolutely conserved, at position, in the six species for which we have data, including the cow (Figure 4c). All S and T residues within the repeat are potential sites for O-linked glycans in human MUC1 (276), and they are presumed to be the major glycosylation sites in other species. P residues are common in the vicinity of O-glycosylation sites, and they may be integral to sequons for O-linked glycosylation (377, 430). Additionally in MUC1, P residues ensure that the polypeptide repeat sequence has little secondary structure. Interestingly, there is a sequon for N-linked carbohydrate (NXS/T, where X is any amino acid except P) within the tandem repeat of bovine MUC1 (Figure 4c). Five such sites are present in the degenerate repeats of the mouse mucin (373). Whether these sites are glycosylated in the mammary mucins in vivo is not known. In toto, the tandem repeat can be viewed as a modular unit for addition of large quantities of covalently-linked carbohydrate to an extended polypeptide backbone in the exoplasmic domain.

In cow, human, and gibbon, the repetitive domains comprise exact repeats that are reiterated a variable number of times (as discussed above, they are bona fide VNTR). However, in mouse and other rodents, the number of repeats is fixed (analyzed in inbred laboratory strains or feral species), and the repeat is not strictly conserved (304, 373). Thus, the repetitive domain in the mouse gene is not polymorphic. Rabbit MUC1 contains 17 degenerate repeats in inbred strains (161). However, in two wild species (jackrabbit and cottontail) a larger repetitive domain was identified by Southern blot, indicating some degree of polymorphism within the order (161).

Posttranslational Modifications

MUC1 has been purified from the MFGM of cow (78, 367), goat (76), human (142, 298, 359, 360), and guinea pig (201), and the preparations from cow, human, and guinea pig have been analyzed for content of amino acids and sugars. As may be expected, the most abundant amino acids are those present within the tandem repeat, which accounts for 45 to over 80% by weight of the peptide chain of individual molecules of human MUC1. The most common residues in human MUC1 are S, T, P, G, and A (359), and similar trends are seen in the bovine (367) and guinea pig mucins (201).

Bovine MUC1 contains an estimated 50% (wt/wt) carbohydrate, which comprises fucose, galactose, mannose, N-acetylgalactosamine, N-acetylgalactosamine, and sialic acid (367). Because of the large content of sialic acid, the mucin has extremely acidic isoelectric points, and following isoelectric focusing, it is recovered in fractions with pH values below 5.0 (262). The bulk of the sugar associated with bovine MUC1 is presumed to be covalently linked to S and T residues within the repetitive domain and flanking regions. At least one site for N-linked glycans is present in the region between the repetitive domain and the membrane anchor in addition to the site within the tandem repeat discussed above (Figure 4b, c). Because bovine MUC1 contains mannose (367), either or both of these sites may be glycosylated. Five sites outside of the tandem repeat are conserved in primate MUC1 [(122, 243, 372, 424); Figure 4b], and there is evidence that human MUC1 contains N-linked glycans, at least in breast cancer cell lines (162, 245).

None of the glycans associated with bovine MUC1 have been structurally characterized. Extensive analysis of the O-linked glycans of human MUC1 has identified linear and branched poly N-acetyllactosamine sequences linked to the polypeptide backbone via N-acetylgalactosamine residues (141, 143). Considerable heterogeneity exists in the number of lactosamine repeats, degree of branching, and presence or absence of terminal sugars, such as sialic acid and fucose. A full discussion of these data and the aberrant glycosylation of MUC1 in mammary carcinomas is beyond the scope of this review [see, e.g., (59, 141, 248, 275)].

We are unaware of any studies documenting the phosphorylation or lipidation of bovine MUC1. As indicated above, human MUC1 is phosphorylated in cultured cells on several S and Y residues (241, 437). Whether MUC1 in human milk is phosphorylated has not been reported.

XDH/XO

XDH/XO is the major component of a prominent Coomassie blue-positive band of protein migrating with an apparent Mr of 155,000 in 10% SDS-polyacrylamide gels (Figure 1a, lane 1; Figure 1c, lane 5). This band was
denoted by various numbers [e.g., (231, 258); Table 1) before its identity was recognized (249, 263, 416). XDH/XO is the most abundant protein with a known enzyme activity in washed milk-fat globules, comprising some 20% of globule-associated protein (263). Approximately half of the enzyme is released into the MFGM supernatant during phase inversion, and the remainder is recoverable with the MFGM by centrifugation (57, 62, 262, 263). At least 60% of the XDH/XO associated with MFGM can be removed by washing the membrane with solutions containing nonionic detergents or high concentrations of salt (58, 62). The residual enzyme remains firmly bound to membranes. These dual soluble-insoluble properties are a reflection of the peripheral nature of XDH/XO in secreted milk-fat globules.

XDH/XO is a molybdenum-containing redox enzyme of considerable complexity, which has been widely studied. Many excellent reviews have been published covering enzymology, protein structure, and potential enzyme-related pathologies (56, 89, 164, 165, 166, 423). The proceedings of a two-day symposium on XDH/XO discussing many of these issues was published recently as a useful compendium of short reviews (149).

Bovine XDH/XO comprises a homodimer of Mr approximately 300,000 (90, 416). Analysis of the amino acid sequence derived from cloned cDNA predicts that each monomer comprises 1332 amino acid residues with an Mr of 146,600 and isoelectric point of 7.7 (39). Each monomer contains a molybdopterin cofactor, two Fe2/S2 clusters, and one molecule of FAD (252) (Figure 5a, b). The monomers are assumed to be catalytically independent of each other (discussed in 56, 89). XDH/XO oxidizes a wide variety of purines and other compounds (56) by a reaction in which oxygen from water is incorporated into the substrate, and the released reducing equivalents are transferred to redox centers within the protein. The reaction can be formally summarized as

\[ R-H + H_2O \rightarrow R-OH + 2H^+ + 2e^- \]

The mammalian enzyme exists as both a dehydrogenase (D-form) and an oxidase (O-form) (167, 381). In the D-form, reducing equivalents are preferentially transferred to NAD\(^+\) with the formation of NADH + H\(^+\). Oxygen can also function as electron acceptor for the D-form in the absence of NAD\(^+\) (146, 179). In the O-form, reducing equivalents are primarily transferred to molecular oxygen with the formation of H\(_2\)O\(_2\) and superoxide radicals (O\(_2^−\)) (56, 264). Purines interact with the enzyme at the molybdopterin site and NAD\(^+\) and oxygen at the flavin site (Figure 5b). In addition,
the enzyme can act as an NADH oxidase by interaction of substrate through the flavin site (56, 148, 252).

The enzyme is predominantly expressed as a dehydrogenase in tissues and is converted to the O-form, either by oxidation of specific cysteine residues or by limited proteolysis (99, 167, 289, 381). Typically, the percentage of oxidase rises during purification of the enzyme because of the in vitro formation of disulfide bonds on exposure to air (99, 167). These O-forms can be reversibly converted to the D-form by addition of thiol-reducing agents (33). In contrast, proteolytic cleavage irreversibly converts the enzyme to the O-form (33). XDH/XO is estimated to be approximately 50 to 75% O-form in freshly collected milk (179, 279). This relatively high level of endogenous oxidase in milk may be due to the oxidation of sulfhydryl groups in the D-form by sulfhydryl oxidase associated with MFGM (83).

Various inactive forms of the enzyme have been described including a demolybdo form lacking the Mo atoms (151) and a desulfo form (251) in which an essential sulfido group at the molybdopterin active site (Mo=S) is replaced by an oxo group (Mo=O) (92). The amount of demolybdo enzyme secreted into milk is presumed to be inversely correlated with the availability of dietary Mo (151, 364). The desulfo form may be a physiologically important inactive form of the enzyme (60, 186), but it is also artifactually generated during preparation or storage (reviewed in (56)). The purified milk enzyme is often more than 60% nonfunctional because of the presence of these two inactive forms [approximately 30 to 40% demolybdo and 20 to 30% desulfo (148)]. Active enzyme can be conveniently separated from the inactive desulfo form by chromatography on folate/Sepharose 4B (290).

Despite a wealth of information on the molecular constitution and kinetic parameters of XDH/XO, the full range of its physiological functions remains uncertain (56, 148, 252). In uricoteric organisms, the enzyme plays a key role in the terminal steps of purine metabolism, catalyzing the oxidation of hypoxanthine to xanthine, and of xanthine to uric acid (305, 364). However, in mammals, the widespread distribution of the enzyme in cells and tissues (63, 193) and the large quantities of XDH/XO associated with milk-fat droplets in many species (254) imply functions beyond the realms of purine metabolism. Furthermore, in many human tissues (305, 364) and milk (1, 60), XDH/XO has relatively low intrinsic activities toward purines compared with the enzyme in other mammals. XDH/XO in human milk may be 98 to 99% nonfunctional as a purine oxidase, at least partly because of the presence of inactive desulfo and demolybdo forms (1, 60). Specific activities in human milk appear to vary widely during the first few weeks of lactation, suggesting temporal control of the enzyme, possibly by interconversion between active sulfo and inactive desulfo forms (60). However, despite such variation in the ability to oxidize purines, human XDH/XO retains a significant NADH oxidase activity, mediated through the flavin site, which is comparable with that of the bovine milk enzyme (1).

There is increasing evidence that XDH/XO in tissues may function as a signaling molecule, serving as a source of reactive oxygen species including H2O2, O2−, ONOO−, and NO (68, 148, 252, 404, 433) that regulate downstream targets including transcription factors. Reactive oxygen species are potent regulators of gene expression and can modulate the immune response (reviewed in (221, 303)). XDH/XO has also been implicated in inflammatory processes, because expression of the XDH/XO gene, protein, and enzyme activity is induced by cytokines and steroids (111, 235, 302, 324). In this capacity, XDH/XO may generate reactive oxygen species that either cause tissue damage and exacerbate the inflammatory response or induce expression of genes encoding, for example, adhesive proteins, cell receptors, and components of the immune system that modulate tissue healing. Reactive oxygen species produced by XDH/XO have been implicated in ischaemia/reperfusion injury, which occurs in cardiac muscle following heart attack and during the in vitro maintenance of organs for transplant (reviewed in references (68, 364)).

The large quantities of XDH/XO in the mammary gland and the MFGM deserve special comment. Expression of the XDH/XO gene increases during late pregnancy and is maximal following parturition (235). Levels of XDH/XO activity and protein in lactating bovine mammary tissue are among the highest of any organ in the body, including liver (63, 305). An exclusive role in purine metabolism is contraindicated by the low levels of uric acid (generally <50 μM) in milk (198). XDH/XO may play a structural and functional role in the formation of the MFGM by binding to the cytoplasmic domain of BTN (188), thus forming a protein complex that interacts with lipid droplets at the apical cell surface (260; see section on BTN below). In addition, XDH/XO may function as an antibacterial component in milk by providing a source of H2O2 for lactoperoxidase (44). In such a capacity, the enzyme may play immunoprotective roles in the gut of the suckling neonate (cited in reference (302)) because a substantial amount of XDH/XO activity may survive passage through the stomach into the small intestine (434). However, the identity of the oxidizable substrates necessary for such functions in milk and the intestinal lumen are unclear.

A confounding paradox in any discussion of XDH/XO function is the occurrence of individuals with hereditary xanthinuria that lack XDH/XO protein (184, 364). In
three patients, this condition was recently traced to nonsense or frameshift mutations in the XDH/XO gene (184). Many such individuals are asymptomatic, apart from obvious metabolic disturbances in purine metabolism, including, for example, low uric acid and high xanthine levels in the serum, renal problems, and the occurrence of xanthine stones (364). Thus, if XDH/XO has additional functions, apart from the metabolism of purines, there must be physiological mechanisms that can compensate for the lack of XDH/XO protein.

Sequence Analysis

Protein sequence data, largely derived from cloned cDNA are available for the eukaryotic enzyme from bovine milk and liver (39, 399), human liver and composite human sources (185, 340), rat liver (6), mouse liver (398), chicken liver (341), Drosophila sp. (87, 219, 331), Bombyx mori (428), and Caliphora vicina (172). Complete structures for the human (425), mouse (79), and Drosophila (87, 219, 238, 331) genes are available; the 5′-region of the rat gene has been sequenced (82, 84), and a putative XDH/XO gene has been identified in the Caenorhabditis elegans genome (73). In addition, the X-ray crystallographic structure of the closely related aldehyde oxidase from Desulphovibrio gigas (401) has been completed (333). This information has provided important insights into the probable structure and mechanism of eukaryotic molybdyl hydroxylases, including XDH/XO (178).

The sequences of mammalian XDH/XO are very similar across species (Figure 5c) with amino acid identity approaching 90% between the bovine, rat, mouse, and human enzymes. Chicken and Drosophila XDH/XO are, respectively, 70 and 53% identical to the mammalian enzymes (167). Limited proteolysis with trypsin or subtilisin cleaves XDH/XO into three major fragments of approximate Mr, 20,000, 40,000, and 85,000 that remain associated as a complex, unless chaotropic agents or detergents are added [(6, 39, 341); see Figure 5c for specific cleavage sites for the bovine, rat, and chicken enzymes]. These fragments essentially define three major domains of the protein (Figure 5a, b). The 20,000-Da fragment contains the amino terminus and two Fe₂S₂ centers complexed to the protein through two sets of four cysteine residues each (6, 165, 167, 341). The central 40,000-Da fragment contains a potential NAD⁺ binding site in the D-form (6, 167, 288) and probably binds the flavin cofactor (165, 167). The C-terminal 85,000-Da fragment contains binding sites for the molybdopterin cofactor (165, 167, 333, 341, 423).

Cysteine residues in bovine XDH/XO that may bind to the Fe₂S₂ clusters and residues, which may interact with the molybdoptereryl moiety, can be tentatively identified by comparison with the structurally characterized aldehyde oxidase from D. gigas (333, 401) (Figure 5c). This enzyme is 39.7 and 29.6% identical to the N- and C-terminal portions of bovine XDH/XO, respectively. Cysteine residues C474-3, C524-8, C555-1, and C77-3 (39) may bind the first Fe₂S₂ cluster, and residues C118-1, C121-1, C155-1, and C157-1 may bind to the second (333). These residues are fully conserved, at position, in all species analyzed (astersisks, Figure 5c). Five other cysteine residues in the N-terminal region are less well conserved [(165, 167); open squares, Figure 5c]. Possible contact sites between molybdopterin and residues within the C-terminal domain can also be tentatively identified (333) (arrowheads, Figure 5c). However in aldehyde oxidase, the molybdopterin cofactor, is a cytosine dinucleotide, and additional contact sites between the pyrimidine base, ribose, and aldehyde oxidase are less conserved in XDH/XO (333).

The exact positions of the flavin and NAD⁺-binding sites remain uncertain, because aldehyde oxidase lacks a flavin-binding domain. Bovine XDH/XO contains the conserved motif GXGXXG, which is involved in nucleotide binding (419) at two positions: residues G464-2 to G51-47 and G847-780 to G852-800 (39). However, neither of these sites correspond to the putative NAD⁺-binding site in the D-form of the enzyme at residues F419-889 to R424-9, identified by comparison with a chemically labelled site in the chicken liver enzyme [(288); circle, Figure 5c]. Furthermore, residues G464-2 to G51-47 overlap with the first potential Fe₂S₂ binding site, and residues G847-780 to G852-800 may interact with the molybdopterin cofactor.

Posttranslational Modifications

Posttranslational processing reactions have not been extensively studied. XDH/XO does not have a recognizable N-terminal hydrophobic leader sequence, and it is presumed to be synthesized in the cytoplasm on free ribosomes. The enzyme is not known to be glycosylated (117, 193), and an earlier report that a serine residue near the Mo center is phosphorylated (96) has been discounted (199). One study has been published (213), indicating that bovine milk XDH/XO is acylated with medium and long-chain saturated and unsaturated fatty acids.

Proteolytically cleaved fragments of XDH/XO can be routinely identified as minor components of bovine MFGM, especially two species with estimated Mr of approximately 130,000 to 140,000 and 90,000 (193, 261). The sizes of these fragments suggest that they are generated by limited cleavage between the three major domains of XDH/XO, although this has not been verified by sequencing. Whether this limited proteolytic
breakdown occurs in situ or in vitro during the preparation of MFGM fractions is uncertain. Similarly sized fragments and additional peptides of approximately Mr 60,000, 40,000, and 20,000 may accumulate in purified preparations of the enzyme (249, 278, 383). Perhaps because of acylated residues, or other unidentified modifications, the major 150,000-Da form of the bovine milk enzyme does not focus as a single species during isoelectric focusing (193, 253). Bovine XDH/XO is resolved into several bands of enzyme by either analytical isoelectric focusing in polyacrylamide gels (253, 266) or by preparative electofocusing in flatbeds of Sephadex (262, 383). The most reliable estimates for the isoelectric points of the milk enzyme range from 6.9 to 7.6 (262, 383), in reasonable agreement with a calculated value of 7.7 from the amino acid sequence derived from cloned cDNA (39).

Polymorphism of the XDH/XO gene has not been extensively studied. An early report identified two alleles based on activity measurements of milk XDH/XO in 92 Guernseys and 39 Holsteins (435). “Low” and “high activity” alleles were described that were additive and displayed no dominance. Whether these forms relate to the inactive and active forms of XDH/XO identified in bovine milk was not established and deserves further study, especially now that the cDNA sequence for bovine XDH/XO is available. Possible polymorphism within exons of the bovine gene is suggested by differences between the cDNA cloned from bovine liver (399) and mammary gland libraries (39). Eleven substitutions, one insertion, and one deletion were noted between the derived amino acid sequences (399).

PAS III

PAS III is a poorly characterized glycoprotein of bovine MFGM and the supernatant fraction, which is resolved in SDS-polyacrylamide gels as a diffuse band of Mr 76,000 to 78,000 [(134, 136); Figure 1a, lane 1]. Because of its large content of carbohydrate [approximately 24% (wt/wt) neutral sugars; (38)] the protein stains strongly with the PAS reagent (Figure 1b, lane 3) and modified silver stain (not shown). CD36 comprises 5% or less of the total protein associated with bovine MFGM (136, 262), and because it is an integral protein (133), most is recovered from the buttermilk in the MFGM fraction by centrifugation (Figure 1a, lane 1; Figure 1b, lane 3). Previous names for the bovine MFGM form of CD36 include PAS IV (134) or Band 3/IV (262) (Table 1). Because this protein has now been unambiguously identified as CD36 (38, 136), the Committee recommends that earlier names based on staining characteristics or electrophoretic mobility be discontinued.

CD36

CD36 can be identified in SDS-polyacrylamide gels as a minor band of Coomassie-blue-positive protein of apparent Mr 64,000 to 68,000 (134, 136; Figure 1a, lane 1). Because of its large content of carbohydrate approximately 24% (wt/wt) neutral sugars; (38)] the protein stains strongly with the PAS reagent (Figure 1a, lane 1) and modified silver stain (not shown). CD36 comprises 5% or less of the total protein associated with bovine MFGM (136, 262), and because it is an integral protein (133), most is recovered from the buttermilk in the MFGM fraction by centrifugation (Figure 1a, lane 1; Figure 1b, lane 3). Previous names for the bovine MFGM form of CD36 include PAS IV (134) or Band 3/IV (262) (Table 1). Because this protein has now been unambiguously identified as CD36 (38, 136), the Committee recommends that earlier names based on staining characteristics or electrophoretic mobility be discontinued.

CD36 was first purified from bovine MFGM as the integral protein PAS IV (134) and was identified by comparing the N-terminal sequence with authentic CD36 from bovine heart or human platelets (135, 136). Human CD36 was originally defined with monoclonal antibodies, OKM5 and OKM8, as a differentiation antigen (cluster of differentiation, CD) on monocytes and platelets (387). Further work with a range of monoclonal and polyclonal antibodies established that CD36 is widely expressed on monocytes and macrophages, erythroid precursors, platelets (20, 222, 230, 387, 389), capillary endothelial cells (134, 230), adipocytes (3), epithelial cells from mammary gland and lung (132, 134), and various cell lines (133, 389).

CD36 has been implicated in diverse functions within and outside the hemopoietic and vascular systems [reviewed in (19, 133)]. As a receptor for collagen (21, 388) and thrombospondin (20, 239, 240, 363), it may function in platelet activation and aggregation (5, 293, 388), intercellular adhesion (362), and the thrombospondin-mediated inhibition of angiogenesis (98). In patients
with malaria, CD36 expressed on endothelial cells, monocytes, and platelets binds to specific proteins (30, 292) on the surface of erythrocytes infected with the trypanosome Plasmodium falciparum (297). As a consequence the parasite becomes sequestered within peripheral tissues and evades destruction in the spleen, thus contributing to progression of the disease [reviewed in reference (133)].

CD36 may also act as a scavenger receptor by binding to apoptotic cells and cell fragments, thus precipitating their elimination by phagocytosis. Examples include the removal of apoptotic neutrophils by macrophages following acute inflammation (281, 345) and phagocytosis by retinal pigment epithelial cells of rod outer segments, constitutively shed from photoreceptor cells (337). CD36 may recognize apoptotic cells by binding to anionic phospholipid (phosphatidylserine) that is abnormally exposed on the rearranged exoplasmic face of necrotic cell surfaces (330, 336, 385).

A related scavenger function lies in the ability of CD36 expressed on the surface of macrophages to bind oxidized low-density lipoproteins and to eliminate them from the systemic circulation by endocytosis (108, 112, 140, 378). Human CD36 also binds native lipoproteins (74), although binding may not necessarily be linked to lipid uptake (88, 139). There is even evidence that CD36 binds long-chain fatty acids and acts as a constitutive fatty acid transporter [(2, 24, 112, 183); however, see also reference (409)].

Despite this functional versatility, it is not clear why CD36 is expressed on the surface of mammary epithelial cells and the MFGM. Because thrombospondin is secreted into colostrum and milk (97), CD36 may act as an epithelial thrombospondin receptor, which functions in the transcytosis of thrombospondin across mammary secretory cells (136). Alternatively, or additionally, the mammary form of CD36 may function as a transporter of long-chain fatty acids, although it is not intuitively obvious in this case why the protein is concentrated on apical rather than basal/lateral plasma membranes.

Sequence Analysis

Complete amino acid sequences are currently available from cloned cDNA encoding CD36 from bovine mammary gland (38), human placenta and erythroid leukemia cells (297, 393), mouse macrophages (108), rat adipocytes (3), and CHO cells (GenBank™ Accession No. U42430). The human gene has been sequenced (17, 18) and several CD36-like genes have been identified in the C. elegans genome (73). Spliced variants of the human transcript have been characterized that differ in either the 5′- (393), or 3′- (291) UTR. In one variant, exons 4 and 5 within the open reading frame are absent because of exon skipping (390). Alternate spliced variants of mRNA for bovine CD36 which differ within the 5′ UTR have also been identified from cloned cDNAs (38).

Bovine CD36 comprises 472 amino acids, including the initiator methionine, which is removed in the mature protein (38). Unprocessed CD36 has a predicted Mr of 52,926 and isoelectric point of 7.93. Similarities between full-length CD36 amino acid sequences from different species are very high. Human, mouse, rat, and hamster CD36 are from 83 to 87% identical to the bovine protein.

A closely related group of proteins has been identified by molecular cloning techniques including rat (412) and human (118) lysosomal protein LIMP II, hamster and rat scavenger receptor SR-BI (4, 129; GenBank™ Accession No. AF071495) and related human homolog CLA-1 (4, 75, 277), Smmp-1 protein from the silk moth (332), and two proteins from Drosophila, emp (150), and croquemort (114, 115). Sequence similarities and key structural features are conserved throughout the membrane-spanning and exoplasmic domains. Human and rat LIMP II are approximately 34% identical to human CD36 (412). Because of these similarities, all of these proteins are considered to be encoded by an extended superfamily of CD36-related genes (133, 412).

Analysis of the amino acid sequences and membrane topology of CD36 family members reveals some remarkable structural features [(133); Figure 6a, b]. The proteins have contiguous stretches of hydrophobic amino acid residues close to both the N- and C-termini (Figure 6a, b). A mutant form of CD36 in which the C-terminal hydrophilic sequence was deleted remained associated with membranes, suggesting that the N-terminal sequence acts as a transmembrane anchor (391) [however also see reference (318) for contradictory results]. The N-terminal sequence, therefore, appears to function as an uncleaved signal sequence (297) and membrane anchor (136) and directs the peptide chain to the endoplasmic reticulum membrane during synthesis. The C-terminal sequence presumably acts as a “stop-transfer” signal. Because C residues in both the extreme N and C termini (Figure 6b) are palmitoylated (391) by acyltransferases presumed to be of cytoplasmic orientation (350), N- and C-termini face the cytoplasm, and both terminal hydrophobic sequences function as membrane anchors (391). CD36-like proteins are, therefore, uniquely associated with membranes as hairpinlike structures with short cytoplasmic sequences at both N- and C-termini and extensive exoplasmic domains (Figure 6a).

The exoplasmic domains also have several conserved hydrophobic sequences that may form hydrophobic
Figure 6. Structure and sequence comparisons of CD36. (a) Proposed domain structure of CD36 showing the two putative transmembrane domains (TM1, TM2), the exoplasmic (exo) hydrophobic region (HR), and the two short cytoplasmic (cyto) tails. N-linked glycosylation sites in bovine CD36 are shown above the diagram (\(\text{H17006}\)). (b) Comparison of the sequences of bovine, human, mouse, and rat CD36 with rat limp2 protein using the PILEUP program [Wisconsin Package Version 9.1. Genetics Computer Group (GCG), Madison, WI]. The N- and C-terminal membrane anchors, exoplasmic hydrophobic region, and thrombospondin binding motifs, blocks A, B, and C, discussed in the text, are indicated above the bovine sequence. N-linked glycosylation sites (\(\text{\textbullet}\)), conserved C residues (*), and putative \(\beta\)-N-acetylgalactosaminylation signal motif in bovine CD36 (\(\text{\textbullet}\)). GenBank TM/EMBL or Swiss Protein (Sw. Prot.) Accession numbers are as follows: bovine CD36, X91503; human CD36, M24795; mouse CD36, Q08857 (Sw. Prot.); and rat CD36, Q07969 (Sw. Prot.) and rat limp2, P27615 (Sw. Prot.).

Pocket in the folded polypeptide. One such sequence between residues L180 through Y209 (38) may be associated with the exoplasmic face of the lipid bilayer [(133); Figure 6b]. This latter sequence divides the exoplasmic region into two putative domains: an N-terminal domain that has no cysteine residues and contains a conserved thrombospondin binding motif (93) and a proline-rich region between amino acid residues C249 and P418 that contains all six exoplasmic cysteine (133). The thrombospondin recognition motif in the N-terminal domain comprises a protein kinase C consensus sequence GPYTYR (Block A) and two closely spaced binding sequences (Blocks B and C) (Figure 6b). Phosphorylation of the T93 residue within Block A is proposed to inhibit binding of thrombospondin through Blocks B and C (discussed in reference 93)).
The six exoplasmic cysteines within the C-terminal region form three disulfide bonds, which in bovine CD36 have been mapped to residues C249-242 and C317-310, C278-271 and C339-332, and C319-312 and C326-321 (38, 329). Other CD36 proteins are presumed to have a similar pattern of disulfide bonds (329) because all six cysteines are conserved at position (asterisks, Figure 6b), even in the insect proteins Snmp1 and croquemort (114, 329, 332). A somewhat more divergent pattern of cysteines is present in CLA1/SR-B1 receptor and LIMPII (Figure 6b; (329)). Formation of one or more disulfide bonds in the exoplasmic domain appears to be necessary for the export of CD36 from the endoplasmic reticulum and intracellular transport through the Golgi complex to the plasma membrane (138).

One other notable feature of the exoplasmic domain is the large number of glycosylation sites. Bovine CD36 contains eight potential sites for N-linked glycans within the sequon, NX/S/T (arrowheads, Figure 6a; diamonds, Figure 6b). Many of these N residues are also present in other family members, either completely conserved at position or within close proximity (Figure 6b). The large amount of N-linked carbohydrate in CD36-related proteins accounts for their proteolytic breakdown, a property that may protect LIMPII from constitutive degradation in lysosomes (133).

The cytoplasmic domains of CD36 family members comprise two extremely short sequences at the N- and C-termini. Of special note are four cysteine residues, two at either end of the protein that are retained in all characterized mammalian CD36 proteins (asterisks, Figure 6b). As previously discussed, all four residues are acylated in human CD36 (391). The two cysteine residues in the C-terminus are within the context of a motif, CXC, with adjacent basic residues (Figure 6b) that in other proteins binds to tyrosine kinases (354, 405). Binding requires two closely spaced cysteine residues within the kinase itself and may occur by coordination of a metal ion (354, 405). Because CD36 has been identified as a signaling molecule (348, 403), this motif may bind to kinases and function in intracellular signaling cascades. Several protein tyrosine kinases of the src family are associated with CD36 in human platelets, melanoma, erythroleukemia (177), and endothelial (69) cells. However, because src kinases that bind to CD36 do not contain the necessary cysteines (133, 177) and because the cysteines in CD36 may be acylated, association cannot be directly through metal coordination. One possibility is that protein-protein interactions are indirect via linker molecules (177).

Outside of the immediate CD36 protein family, the cytoplasmic sequences of related proteins vary in length and display little identity. The C-terminal extension peptide of rat LIMPII (Figure 6b) functions as a lysosomal targeting signal (411) and explains how CD36 and LIMPII are directed to such different cellular locations, despite their structural similarities.

**Posttranslational Modifications**

Bovine and human CD36 are heavily glycosylated. Carbohydrate accounts for at least 26% by weight of the protein from human platelets (389) and, excluding sialic acid, 24% of CD36 in bovine MFGM (38). Differences in Mr, between the species can be largely accounted for by the variable addition of N-linked glycans (38, 136). Removal of carbohydrate by digestion with endoglycosidase F leads to reduction in Mr of various forms of bovine and human CD36 from a range of 76,000 to 88,000 to a similar sized protein of Mr, 57,000 (136). This latter value is close to the size of approximately 53,000 predicted from the sequenced cDNA (38, 297).

The N-linked glycans associated with CD36 in bovine MFGM have been extensively characterized (38, 280). In one study, CD36 was estimated to contain 6 mol of asparagine-linked carbohydrate/mol of protein (280). However, all eight of the possible glycosylation sites (Figure 6a, b) were shown to contain carbohydrate by analysis of peptides encompassing each site (38). This finding suggests that potentially all positions can be glycosylated, although not necessarily in all molecules.

Glycans associated with CD36 from bovine MFGM comprise high mannose-, hybrid-, and complex-type chains (bi-, tri-, and tetraantennary) [(280), Table 2]. Both the hybrid-type and complex-type chains may be terminally sialylated (280). The high mannose-type chains and many of the complex-type and hybrid-type chains are typical of those found in other N-linked glycoproteins. However, in one study, 28% of the asparagine-linked glycans recovered from bovine CD36 by hydroxynolysis contained GalNAcβ1→4GlcNAc structures (280). This sequence has been identified in other bovine glycoproteins (339, 342), including BTN and PAS 6/7 (280). This sequence has been identified in other bovine glycoproteins (339, 342), including BTN and PAS 6/7 (280). This sequence has been identified in other bovine glycoproteins (339, 342), including BTN and PAS 6/7 (280). This sequence has been identified in other bovine glycoproteins (339, 342), including BTN and PAS 6/7 (280).
hybrid structures with either Galβ1→4GlcNAc or GalNAcβ1→4GlcNAc moieties in the Manα1→3 arm (Table 2). Analysis of the gross sugar composition of peptides containing each of the eight glycosylation sites shows that some asparagine residues are selectively glycosylated with particular types of oligosaccharide (38). Sugar chains associated with residues N210204, N253246, and N423416 (38) contain large amounts of mannose and smaller quantities of either galactose or GalNAc. These chains are likely to be mainly of the high mannose-type with some complex-type structures. GalNAc is associated with residues N8078, N103101, N174171, N327320, and N423416 and will be bound in β1→4 linkage to GlcNAc in complex- or hybrid-type structures (38, 280).

The biochemical basis for such site-selective or site-specific glycosylation is poorly understood, although a recognition motif for the addition of GalNAc to complex chains has been suggested [(365); see below]. The addition of GalNAc in β1→4 linkage to terminal GlcNAc residues in partially processed glycoproteins requires a novel GalNAc transferase, which is presumably located in the Golgi complex together with other terminal glycosyl transferases. For the human glycoprotein hormones, luteinizing hormone and chorionic gonadotrophin, it has been suggested that addition of GalNAc requires the recognition motif, PXR/K, six to nine residues upstream of the glycosylation site in the protein backbone (365). However, there is only one such motif in bovine CD36, PAK, encompassing residues P401394 to K403396 (arrowheads, Figure 6b), which is not close to any of the glycosylation sites with oligosaccharides.

Table 2. CD36-associated N-linked glycans after desialylation.1

<table>
<thead>
<tr>
<th>Structures</th>
<th>Molar ratio (%)</th>
<th>Structures</th>
<th>Molar ratio (%)</th>
<th>Structures</th>
<th>Molar ratio (%)</th>
<th>Structures</th>
<th>Molar ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
</tr>
<tr>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
</tr>
<tr>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
</tr>
<tr>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
<td>Galβ1→4GlcNAc</td>
<td></td>
</tr>
</tbody>
</table>

1Table adapted with permission from Nakata et al. (280). Copyright 1993. American Chemical Society. R = GlcNAcβ1→4 (Fucα1→6)GlcNAcOT; R' = GlcNAcβ1→4GlcNAcOT.
charides shown to contain GalNAc (38). Other β-N-acetylgalactosaminylated glycoproteins have been characterized [including PAS 6/7; (181)] that do not possess the PXR/K motif (100). These are presumably glycosylated by a different class of transferase enzyme (100).

The unusual mannose chains in CD36 and other bovine glycoproteins indicate that the processing of N-linked glycoproteins in bovine tissues may occur by alternate pathways. In addition, this mannose structure occurs in hybrid-type glycans with the uncommon GalNAcβ1→4GlcNAc structure. A number of atypical processing and glycosylation reactions are possible that would require α-mannosidases and GalNAc transferases with unique specificities [discussed in reference (280)].

One other posttranslational modification to CD36 that has been well characterized is the addition of palmitic acid by thioester linkage to cysteine residues in the cytoplasmic termini. For human CD36, site-directed mutagenesis was used to show that all four cysteine residues (two in the N-terminal and two in the C-terminal, Figure 6b) are acylated (391). These acyl chains, close to both membrane anchors, may help to stabilize the association of CD36 with the lipid bilayer. Whether bovine CD36 is acylated in a similar fashion is unknown, although all four cytoplasmic cysteine residues are conserved in both species.

**BTN**

BTN is the most abundant protein in bovine MFGM and has been identified in Coomassie-blue-stained gels as a prominent band of apparent Mr 66,000 to 67,000 [(231, 258, 262); Figure 1a, lane 1; Figure 1c, lane 5]. By densitometric analysis, BTN has been estimated to comprise 34 to 43% of the total MFGM protein in Holstein milk (262, 270) and approximately 20% in Jersey milk (270). The protein was named butyrophilin from the Greek butyros/philos meaning ‘having an affinity for butterfat,’ because the protein is concentrated in milk-fat globules, and it is specifically expressed in lactating mammary tissue (116). Previous names include Band IV/Glycoprotein E (9, 12), Component V/Glycoprotein 6 (231), and Component 12/IV (258) (Table 1). Although the function of BTN is not clear (see below), this name has become widely used for the bovine protein and homologs in other species [e.g., (155, 200, 282, 395)]. The Committee recommends that BTN be adopted as the name for this protein in all species for which sequence data are available and that the use of arabic and roman numerals be discontinued.

Most BTN associated with washed fat globules is recovered in the MFGM pellet when buttermilk fractions are separated by centrifugation (Figure 1a, lane 1; Figure 1c, lane 5). A small fraction remains in the MFGM supernatant, presumably in the form of protein aggregates or associated with lipid fragments (Figure 1a, lane 2; Figure 1c, lane 6). MFGM-associated BTN is firmly membrane bound and resists extraction with chaotrophic agents and detergents (117, 263) including 1% (wt/vol) solutions of SDS (257). BTN can be solubilized in SDS solutions by addition of thiol-reducing agents (231). These properties suggest that BTN is an integral protein and that disulfide bonds play a role in stabilizing its association with the membrane (260). The integral character of BTN has been confirmed by molecular cloning techniques (189) and topological assays (28).

**Sequence Analysis**

The first complete amino acid sequence of BTN was derived from a cDNA cloned from a lactating bovine mammary gland library (189). Gene and cDNA sequences for bovine (95, 189, GenBank™ Accession No Z93323), human (334, 395), and mouse (188, 294) BTN and a cDNA sequence for guinea pig BTN (413) are currently available. In addition, the porcine BTN gene has been partially sequenced (414). Overall similarities in gene structure and cDNA sequence are high across species (95, 189, 294, 334, 395, 413, 414). The derived amino acid sequence for bovine BTN is 84% (395), 71% (294), and 74% (413), similar to human, mouse, and guinea pig BTN, respectively.

Bovine BTN comprises 526 amino acids and has an N-terminal leader sequence of 26 amino acids, which is cleaved during processing in vivo (189). The calculated Mr, for the entire amino acid sequence, including the leader peptide, is 59,262 (257). The proteolytically processed protein without any further modifications has a predicted Mr of 56,460 and an isoelectric point of 4.96 (189). There is a single membrane anchor of 27 contiguous hydrophobic amino acids approximately in the middle of the sequence (Figure 7a, b). Topological analysis has shown that BTN is a Type I membrane glycoprotein with the N-terminal domain facing the exoplasmic space and the C-terminal domain facing the cytoplasm (28). This topology appears to be maintained in secreted milk-fat globules (28), although other possibilities have been suggested [for discussion see (28, 117, 254, 257)].

BTN is a member of the Ig superfamily (120, 325), a large family of adhesive proteins, receptors, and components of the immune system (420). Hallmarks of this family are the presence of Ig-like folds in the exoplasmic domain comprising two sheets of anti-parallel β-strands connected by a disulfide bond. Immunoglobulin domains are classified into variable, intermediate, or con-
stant types, depending on the sequence patterns and numbers of strands. On the basis of sequence comparisons, BTN is predicted to have two such domains, an

intermediate-type (IgI) toward the N-terminus (120, 145), stabilized by a disulfide bond between residues, C51^{100} and C125^{124} (189), and a C1 fold (IgC1) toward

![Figure 7. Structure and sequence comparisons of butyrophilin (BTN). (a) Proposed domain structure of BTN showing the signal sequence (SS) and the two proposed Ig folds (IgI and IgC1) in the exoplasmic (exo) domain, the transmembrane anchor (TM), and the conserved B30.2 domain (B30.2) in the cytoplasmic (cyto) tail. (b) Comparison of the sequences of bovine, human, guinea pig, and mouse BTN using the PILEUP program (Wisconsin Package Version 9.1 Genetics Computer Group (GCG), Madison, WI). The signal sequence and cleavage site (*) conserved C residues in the IgI and IgC1 domains; (♦) N-linked glycosylation sites in the exoplasmic domain; (♀) putative β-N-acetylgalactosaminylation signal motif in bovine BTN. GenBank™/EMBL Accession numbers are as follows: cow, M35551; human, U39576; and mouse, U67065. Guinea pig sequence (413).](image-url)
the membrane anchor, stabilized by a disulfide bond between residues C165\textsuperscript{164} and either C219\textsuperscript{218} or C220\textsuperscript{219} ([25, 246, 395]; asterisks, Figure 7b). The sequence of the IgI domain is 33% identical to a variable Ig fold in the chicken B-G antigens (268, 415), and both Ig folds are 40% identical to the Ig variable and IgC domains of B-G-like proteins in cranes (194) [polymorphic components of the avian major histocompatibility complex (MHC), 210]. The IgI domain of BTN is also 46% identical to a variable Ig fold in bovine myelin oligodendrocyte glycoprotein, a component of the myelin sheath (120, 325). In addition, weak homology has been identified between the IgI and IgC1 domains of BTN and similar domains in the B7.1 (CD80) and B7.2 (CD86) receptors (25, 246) (costimulatory molecules expressed with MHC proteins on antigen-presenting cells). Some of the residues conserved between these proteins are key determinants of the Ig domains, suggesting that BTN and the B7.1/7.2 receptors have similar tertiary structures in their exoplasmic domains (25, 246).

The C-terminal, cytoplasmic domain of BTN is characterized by a highly conserved region known as the B30.2 or rfp domain ([81, 156, 158, 189, 415]; Figure 7a, b). This domain comprises approximately 170 amino acids and accounts for almost 70% of the cytoplasmic tail of BTN immediately distal to the membrane anchor. Computer-assisted structural analysis suggests that the B30.2 region is globular in character (187, 422) and contains up to 15 β-strands folded into two Ig-like domains (352). These intracellular Ig domains are not predicted to be stabilized by disulfide bonds (unlike the Ig folds of the exoplasmic domain) because C residues are not conserved within the B30.2 region, and the redox potential of the cytoplasm typically ensures that the C residues of most cytoplasmic proteins are in the reduced form (352). The extreme C-terminal region of 55 to 58 amino acids in BTN is less similar across species; only 28 residues are completely conserved between bovine, human, mouse, and guinea pig BTN (Figure 7b).

The B30.2 domain is conserved in the C-terminal halves of at least 20 other proteins that can be separated into three distinct groups (156, 158). One group, exemplified by ret finger protein (386), the 52-kDa nuclear antigen A of Sjögren’s syndrome (SSA/Ro) (81), and the MID1 protein (328), consists of a family of zinc finger proteins containing RING finger, B-box motifs, and coiled-coil domains in their N-terminal halves (35, 156, 158, 415). Many of these proteins are located in the nucleus and may interact with other proteins or nucleic acids (35, 36). A second group comprises the putative products of several BTN-like genes, containing BTN-like Ig domains in their N-terminal halves and a single membrane anchor (157, 334, 397). By analogy with other members of the Ig superfamily, these proteins may be expressed on the cell surface and function as receptors or adhesive proteins (145, 420). A third group consists of the α and β subunits of stonustoxin, a toxic protein secreted into the venom of the stonefish (Syndaceja horrida) (125). Both subunits contain B30.2 domains in their C-terminal halves and unique N-terminal regions (156, 158). The B30.2 domain is the only conserved sequence currently known that occurs in proteins with such diverse locations and presumptive functions (i.e., cytoplasmic or nuclear factors, membrane receptors, and secreted toxins).

The remarkable conservation of the B30.2 domain in proteins from fish to higher mammals (156, 158, 244) suggests that it has a universal function, possibly in protein-protein interactions (158, 189) mediated through the two predicted Ig folds. Two potential interactive partners are microtubules (349) and XDH/XO (188). XDH/XO has been shown to bind to the cytoplasmic tail of BTN in in vitro assays (188). However, it is currently unclear whether the XDH/XO-binding sites are wholly within or outside of the B30.2 domain or in both regions. Direct binding of XDH/XO to BTN would explain how XDH/XO becomes concentrated over 150-fold on a protein basis in the MFGM (261), over levels in the cytoplasm (193). Furthermore, formation of a supra-molecular complex among XDH/XO, BTN, and proteins on the surface of cytoplasmic lipid droplets, such as ADPH (153, 154), may be an essential step in the assembly of the MFGM and subsequent expulsion of lipid droplets from the cell [discussed in (260)]. In such a scheme, XDH/XO may be viewed as a linker protein between BTN in the membrane bilayer and proteins on the surface of the lipid droplet [see Figure 5c of (260)].

At least six additional BTN-like genes have been discovered in the human MHC Class I region all within 110 kb of each other and the human BTN gene (157, 334, 397). Several of these genes appear to be expressed widely in tissues, at least at the RNA level (334). All the predicted gene products have IgI-variable-like and IgC1-like folds in the putative exoplasmic N terminal domains but two are truncated in the C-terminal domains and lack a B30.2 region. The six genes are grouped into two clusters of three genes each. Genes within each cluster share approximately 95% identity at the amino acid level. Each cluster is approximately 50% identical to the other and to the BTN gene (157, 334, 397). An expanding number of other BTN-like genes is currently being identified and sequenced in the human and mouse MHC (156). It will be interesting to determine whether bovine homologs are located in the bovine MHC (BoLA) on chromosome 23 (22, 65, 392).

Sequence comparisons have, therefore, revealed a canonical BTN-like protein comprising, in order from the
Posttranslational Modifications

The number of isoelectric variants identified by two-dimensional gel electrophoresis [(28, 116, 128, 155); Figure 2b] indicates extensive modifications to the amino acid sequence of BTN. At least four to six variants of bovine BTN are routinely detected with isoelectric points between pH 5.0 and 5.4 (28, 116, 155). These variants can be resolved whether the MFGM is prepared from the pooled milk of several cows or from the milk of individual cows (I. H. Mather, unpublished observations). The two most well-characterized processing reactions are the removal of the N-terminal leader sequence (discussed above) and the covalent addition of N-linked glycans.

Bovine BTN has three potential N-linked glycosylation sites at N56(52) and N216(215) in the N-terminal exoplasmic domain (diamonds, Figure 7b) and N438(437) in the C-terminal cytoplasmic domain. Only residues N56(55) and N216(215) are glycosylated (344), in agreement with the proposed Type I topology of BTN (28, 189). The glycans associated with N56(55) and N216(215) have been extensively characterized and comprise bi-, tri-, and tetraantennary sugar chains with some of the same unusual structures detected in bovine CD36 (280, 344; Tables 2 and 3). Each N residue is apparently glycosylated in a site-specific manner. Glycans associated with N56(55) contain complex-type sugar chains with the universal structure Gal/1→4GlcNAc and the uncommon structure GalNAcβ1→4GlcNAc (344). In addition, both N56(55) and N216(215) residues are glycosylated with high-mannose-type chains. Residue N216(215) uniquely contains hybrid-type glycans with Manα1→2Manα1→3-Manα1→6 and Gal/1→4GlcNAcβ1→2Manα1→3 structures and complex-type oligosaccharides with Galβ1→4GlcNAc structures. Glycans with the uncommon GalNAcβ1→4GlcNAc structure are restricted to residue N56(55). Interestingly, the proposed recognition motif PXR/K in the protein backbone for β-N-acetylglactosaminyltransferase 3 (365) occurs as the sequence P50(52)CR52(51), close to the N56(55) glycosylation site (arrowheads, Figure 7b). However, the C residue is presumed to form a disulfide bond within the IgI domain. Furthermore, as noted above, this sequon is not present at the expected positions in bovine CD36 (38).

BTN is not known to contain O-linked glycans, and digestion with O-glycanase has no effect on the electrophoretic mobility of BTN in SDS-polyacrylamide gels (77, 408). Other possible modifications to BTN include incorporation of covalently bound phosphate (154, 375) and fatty acids (154, 213). BTN contains many S, T, and Y residues, which could be phosphorylated by specific protein kinases (257), and the protein is phosphorylated in vitro when MFGM is incubated with ATP (154, 375). Currently, it is unclear whether BTN is phosphorylated in vivo. Tightly bound medium- and long-chain fatty acids have been identified in gel-purified samples of bovine BTN, including myristate, palmitate, stearate, and oleate (154, 213, 257).

BTN may be proteolytically broken down, either in situ or during the handling of MFGM in vitro (170). The most commonly identified fragment of BTN in MFGM is about 4 kDa smaller than the intact protein and can be readily distinguished in one-dimensional SDS-polyacrylamide gels (Figure 1a, lane 1, small arrow) and by two-dimensional gel electrophoresis (Figure 2a, a′, b; asterisks). This fragment has an intact N-terminus and lacks a portion of the C-terminus (28). Other less-abundant fragments can be identified by immunoblotting techniques. These fragments have been partially characterized with the use of sequence-specific antibody techniques. These fragments have been partially characterized with the use of sequence-specific antibody techniques. These fragments have been partially characterized with the use of sequence-specific antibody techniques. These fragments have been partially characterized with the use of sequence-specific antibody techniques. These fragments have been partially characterized with the use of sequence-specific antibody techniques.

In addition to proteolytic breakdown, BTN also has a marked propensity to aggregate either in situ or during preparation of membrane samples for SDS-PAGE (282). These aggregates can be identified in SDS-polyacrylamide gels as a series of immunoreactive bands of Mr >150,000 [see, e.g., Figure 3 of reference (28)]. From the sizes of these aggregates, it is apparent that both intact and degraded BTN can aggregate. Western blots of bovine MFGM developed with antibody to BTN can, therefore, present a complex pattern of immunoreactive bands that are of lower and higher Mr than the major 66,000- to 67,000-Da band of intact monomeric protein (28).

Cloning and sequence analysis of the bovine BTN gene [(95); GenBank Accession No. Z93323] and amplification of genomic DNA by PCR has identified a number of polymorphic forms of the bovine gene (208, 392, 432). The polymorphic forms include nucleotide substitutions within introns and in exons encoding both exoplasmic and cytoplasmic domains of the protein. At
least seven base changes lead to four changes in the amino acid sequence derived from the original cloned cDNA ([95, 189, 392]; Table 4). Several RFLP can be identified within both introns and exons (392, 432), and a single-stranded conformation polymorphism in the 3′ untranslated region has been described (208).

ADPH

ADPH is a major constituent of the insoluble fraction remaining after MFGM is extracted with salts and non-ionic detergents (154). This protein was overlooked until recently, because it is relatively insoluble in SDS sample buffers, and it has a very similar M₆ to the most abundant form of PAS 6/7 (154) (Figure 1a, lane 1; Figure 1c, lane 5). Consequently, both proteins comigrate as a band of protein of apparent M₆ 52,000 in 8 to 16% (wt/vol) gradient SDS-polyacrylamide gels. Effective solubilization of ADPH requires extended incubation times at elevated temperatures with relatively high concentrations of SDS (10% wt/vol) (154). Bovine ADPH can be resolved from PAS 6/7 by two-dimensional gel electrophoresis and focuses as two major variants with isoelectric points of 7.5 and 7.8 (153, 154).

Bovine ADPH was identified by sequencing tryptic peptides prepared from protein blotted to a membrane. A high proportion (83%) of the sequenced residues was identical to regions within the sequence of mouse ADRP, derived from a full-length cDNA (154, 197). At the time of initial characterization, ADRP was pre-
Table 4. Summary of polymorphisms within the coding region of the bovine butyrophilin gene. 

<table>
<thead>
<tr>
<th>Nucleotide Position</th>
<th>Amino acid Substitution</th>
<th>Position</th>
<th>Amino acid Substitution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>748 (765) T → A</td>
<td>D → E</td>
<td>231</td>
<td>D → E</td>
<td>95</td>
</tr>
<tr>
<td>1315 (1332) C → G</td>
<td>P → P</td>
<td>420</td>
<td>P → P</td>
<td>392</td>
</tr>
<tr>
<td>330 (1347) C → T</td>
<td>V → V</td>
<td>425</td>
<td>V → V</td>
<td>392</td>
</tr>
<tr>
<td>1461 (1478) A → G</td>
<td>K → R</td>
<td>469</td>
<td>K → R</td>
<td>392</td>
</tr>
<tr>
<td>1502 (1519) G → A</td>
<td>I → R</td>
<td>483</td>
<td>I → R</td>
<td>392</td>
</tr>
<tr>
<td>1615 (1632) A → G</td>
<td>P → P</td>
<td>520</td>
<td>P → P</td>
<td>392</td>
</tr>
<tr>
<td>1621 (1638) C → T</td>
<td>P → P</td>
<td>522</td>
<td>P → P</td>
<td>392</td>
</tr>
</tbody>
</table>

1For RFLP see (432).

2Revised numbering submitted to GenBank™ (Accession No. M35551) on December 18, 1996. Original numbering of Jack and Mather (189) shown in parentheses.

suessed to be specifically expressed in adipocytes at a relatively early stage of differentiation (196, 197). Subsequent work has established that ADRP is associated with lipid droplets in many cultured cell lines including fibroblasts, Leydig cells, MDCK cells, and Hep2 hepatoma cells (55, 153). In tissues, the distribution of ADRP appears to be more restricted, essentially to certain cells that accumulate lipid droplets, including the mammary epithelium (153, 154), the livers of etomoxir-treated rats (379), steroidogenic cells in the adrenal cortex, and Sertoli and Leydig cells in the testis (153). ADRP appears to be maximally expressed during early stages of adipocyte differentiation and is absent from lipid storage droplets in mature 3T3-L1 adipocytes in vitro (55) and in adipose tissue in vivo (153).

Heid et al. (153) have recently suggested that ADRP be renamed ADPH to emphasize its association with lipid droplets in many cell types. This new and more generic name is also recommended for general use by the Committee.

Sequence Analysis

Complete amino acid sequences of ADPH are available from cloned mouse and bovine cDNA (197, 285), the characterized murine gene (106), and a compilation of human EST (153). Sequence similarities are very high across species; bovine ADPH is 87 and 80% identical to the human and mouse proteins, respectively (285) (Figure 8). An anti-peptide antibody to the N-terminal first 26 residues of mouse ADPH cross-reacts with the presumptive homolog in rat MFGM (154).

The only known proteins similar to ADPH are proteins in the perilipin family (131) and TIP47, a protein that functions in membrane trafficking between endosomes and the trans-Golgi network (101) [also cloned as protein variants PP17a and b from human placenta, (400)]. Perilipins are phosphorylated proteins encoded by a
single gene that are concentrated on the surface of cytoplasmic lipid droplets in adipocytes (45, 130) and steroidogenic cells (54, 351) but are absent from mammary epithelial cells (45). Like ADPH, the perilipins are associated with the surface of lipid droplets in the cytoplasm. Recent evidence suggests that tumor necrosis factor-α stimulates lipolysis in adipocytes by decreasing the amount of perilipin associated with the lipid droplet surface (370). Thus, perilipins may stabilize lipid droplets from enzymatic breakdown in the absence of physiological agonists (45, 370).

TIP47 binds to the cytoplasmic domain of both the cation-dependent and -independent mannose-6-phosphate receptors (101). These receptors function in the sorting of lysosomal enzymes in the secretory pathway (232). By use of cell-free in vitro assays, Diaz and Pfeffer (101) showed that TIP47 is required for the retrieval of mannose-6-phosphate receptors from endosomes to the trans-Golgi network and that binding is mediated through a FW motif in the cytoplasmic tail of the receptor. Whether TIP47 binds to the surface of lipid droplets, like ADPH and perilipin, has not been reported, although dual functions involving binding to both protein and lipid are possible ([285]; and see below).

The function of ADPH remains unknown. Possible roles in the cellular uptake of long-chain fatty acids (119) and in the accretion and transport of intracellular lipid have been suggested (55, 119, 153, 154). The large amount of ADPH in secreted milk-fat droplets may be fortuitous. On the other hand, interactions between ADPH on the lipid droplet surface and proteins in the apical cytoplasm and plasma membrane, including XDH/XO and BTN, may be required for the formation of MFGM and the budding and secretion of lipid droplets from the cell [(153, 154); discussed in reference (260)].

Surprisingly, despite the close association of ADPH and perilipin with lipid droplets, neither protein has extensive contiguous sequences of hydrophobic amino acids (131, 197). Approximately half of the first 50 residues in the N-termini are nonpolar, which could conceivably form a hydrophobic domain for interaction with neutral lipids [(154); Figure 8]. These amino acids lie within the most similar regions of ADPH and perilipin, comprising the first 110 residues of the N-termini. Nielsen et al. (285) have suggested that conserved residues within two regions (I and II; Figure 8) may form the hydrophobic faces of amphipathic α-helices (L5643, C6047, and M6249 of region I and A9653, L10057, I10360, and L10764 of region II; arrowheads, Figure 8). Hydrophobic residues are conserved at these positions in all characterized ADPH, perilipins, and TIP47 (Figure 8). Furthermore, these regions are both flanked by conserved P residues (asterisks, Figure 8), which may introduce bends in the protein and position the amphipathic α-helices on the protein exterior (285). Thus all proteins within the ADPH family may contain conserved hydrophobic binding domains in their N-termini, which could mediate associations with lipid droplets or in the case of TIP47, the cytoplasmic faces of endomembranes (285).

Posttranslational Modifications

Posttranslational modifications and processing reactions in ADPH have not been extensively studied. ADPH does not contain an N-terminal signal sequence, and it is apparently synthesized on free polysomes in adipocytes [cited in reference (55)]. However, in one study, ADPH was detected in endoplasmic reticulum membrane fractions (154) of lactating mammary glands as well as in cytoplasmic lipid droplets and the MFGM. The microsomal form of the protein (Mr 47,000) was approximately 5 kDa smaller than the form associated with lipid droplets (Mr 52,000) (154), suggesting that ADPH is modified following synthesis. The reason for this discrepancy in the initial location of ADPH in adipocytes and mammary tissue is not clear. Apart from the possible existence of differently sized, intracellular precursors, the secreted MFGM forms of bovine and human ADPH exist as two major isoelectric variants [bovine at pl 7.5 and 7.8 (153, 154) and human at pl 7.0 and 7.2 (153)], also suggesting posttranslational modifications.

ADPH does not appear to be glycosylated, despite the presence of two or three consensus sequences for N-linked glycosylation in the bovine, human, and mouse proteins (153, 154, 197, 285). This finding is consistent with the absence of a signal sequence and location of ADPH in the cytoplasmic compartment. Attempts to show that ADPH is phosphorylated in vivo by injecting [32P]phosphate into a rat or by incubating the radiotracer with isolated bovine mammary acini were unsuccessful (154). However ADPH in bovine MFGM was phosphorylated in vitro when membrane was incubated with [γ-32P] ATP (154). The possible functional significance of this observation is unclear.

ADPH contains tightly bound fatty acid (5 to 6 mol of fatty acid/mol of protein), including myristate, palmitate, stearate, and oleate (154). These acids can be released by treatment with 0.5 M methanolic KOH, suggesting that they are bound to the protein by ester bonds. Fatty acylation of ADPH would markedly increase the hydrophobicity of the protein and possibly aid interaction with neutral lipid droplets in vivo.

PAS 6/7

Identification and resolution of PAS 6/7, the bovine homolog of mouse MFG-E8 (382), has a long and com-
plex history. Two major bands of bovine MFGM protein, which stain with Coomassie blue and the PAS reagent, with literature values for M, ranging from 43,000 to 59,000 (15, 31, 181, 207, 212, 224, 231, 258, 262, 327, 356) are resolved by SDS-PAGE (Figure 1a, lane 1; Figure 1b, lane 3; Figure 1c, lane 5). These proteins were characterized as peripheral components of MFGM because they can be displaced from washed fat globules and isolated membrane by treatment with aqueous solutions containing high concentrations of MgCl₂, other salts, or chaotropic agents [e.g., (131, 181, 207, 224, 258, 357)]. In addition, both proteins are major components of the soluble buttermilk supernatant fraction [e.g., (206); Figure 2a, lane 2; Figure 1b, lane 4; Figure 1c, lane 6] and occur in soluble form in skim milk (72, 256, 327). The proteins were identified by various Arabic or Roman numerals to denote their relative positions in SDS gels and reactivity with dyes (Table 1). Although it was recognized almost 20 yr ago that the proteins are immunologically similar (262), their exact relationship was uncertain until cDNA were recently cloned and sequenced (14, 181). This uncertainty was due in part to their peripheral nature, the resolution of several isolectric variants [(15, 224, 253, 262, 356); Figure 2a, a'], and similarity in size with ADPH (153, 154). Furthermore, comparison of bovine MFGM proteins with those of other species showed only limited similarities in this region of the electrophoretic protein profiles (254). It is now clear that PAS 6/7-like proteins are major components of the MFGM of several species and that most of the heterogeneity observed, including size difference, is due to posttranslational modification of a single protein.

Over the years, many names have been given to this protein to describe electrophoretic mobility, staining characteristics, or molecular size (Table 1). In addition, the human homolog was recently named lactadherin (394) to describe its origin in milk and its adhesive properties. PAS 6/7-associated glycans may bind to rotaviruses and protect the neonatal gut from infection (284, 323, 429), and an RGD motif may mediate binding to integrins [7, 91, 181, 394]; see below]. However, recent cloning of a rat homolog of PAS 6/7 from a brain cDNA library suggests that the protein may be the enzyme O-acetyl ganglioside synthase (295). This enzyme catalyzes the acetylation of the C9 hydroxyl group of terminal α,2,8-linked sialic acid in GD3 ganglioside. Whether the MFGM forms of PAS 6/7 are functional as acetyl transferases is currently unknown.

Because of these considerations, the Committee recommends that the term PAS 6/7 be used for this protein until its physiological and enzymic functions are clearly delineated. If the MFGM forms of PAS 6/7 prove to be an acetyl transferase, the official Enzyme Commission name for this protein can be adopted.

**Sequence Analysis**

Complete amino acid sequences, derived from cloned cDNA are available for bovine (14, 181), human (91, 237), porcine (110), murine (382), and rat (295) proteins. Partial sequences of the guinea-pig homolog GP-55 have also been obtained by Edman degradation of selected tryptic peptides (255). Approximately 9 kb of the mouse MFG-E8 gene have been cloned and partially sequenced (301). Similarities are high across species; bovine PAS 6/7 is from 65 to 85% identical to the derived amino acid sequences from the other cloned cDNA. PAS 6/7 proteins are characterized by N-terminal EGF-like regions and C-terminal tandem repeats that are similar to the C1 and C2 domains of the blood clotting proteins, factors V and VIII (Figure 9a). One EGF domain in each protein contains the adhesive sequence, RGD, which binds to integrin receptors on cell surfaces (7, 91, 181, 394). The C2-like, C-terminal domains contain putative phospholipid-binding motifs, comprising amphiphatic α-helices (7, 181). Topological analysis of the bovine (258) and human (321) proteins indicates that they are exoplasmic, peripheral proteins, bound to anionic phospholipids in the membrane via the C-terminus.

The structure, topology, and tissue distribution of PAS 6/7-like proteins suggests that they function as adhesive proteins. PAS 6/7 is widely distributed in tissues (7, 72), and the RGD sequence may mediate adhesive interactions by binding to integrin receptors on adjacent cells (7, 182). PAS 6/7 is also a component of sperm plasma membrane [p47; (109, 110)] and binds to zona pellucida glycoproteins on the surface of oocytes, thus suggesting a possible adhesive role in fertilization (109, 110). Why such a potentially adhesive protein is expressed on the exposed apical surfaces of the mammary epithelium and secreted milk-fat globules is unclear. As indicated above, human PAS 6/7 (lactadherin) binds to rotaviruses and may protect the neonatal gastrointestinal tract from infection (284, 323, 429). Thus, the prime function of PAS 6/7 in the mammary gland may lie in maternal or neonatal immunity and not in lactation physiology, per se.

Bovine PAS 6/7 comprises 427 amino acid residues, including an N-terminal leader peptide of 18 amino acids, which is cleaved during intracellular processing. The predicted Mₚ of the unprocessed protein is approximately 47,500. Cleavage of the signal sequence yields a protein of approximately Mₚ 45,600 with an isoelectric point of 7.0. Immediately following the signal sequence, the first 89 amino acid residues of bovine PAS 6/7 com-
Figure 9. Structure and sequence comparisons of periodic acid Schiff 6/7 (PAS 6/7). (a) Proposed domain structure of bovine PAS 6/7 showing the signal sequence (SS) and two epidermal growth factor (EGF)-like domains (EGF1, EGF2) in the N-terminus followed by tandem C1/C2-like repeats and an amphipathic helix (AH) at the C-terminus. The region of the C1 repeat missing in the splice variant (SV), discussed in the text, is outlined by dotted lines. N- or O-linked glycosylation sites in bovine PAS 6 only (♦), O-linked glycosylation site in PAS 7 only (♦), N-linked glycosylation site in both PAS 6 and 7 (181) (half black/half white triangle), and position of RGD sequence (v). (b) Comparison of the sequences of bovine PAS 6/7 with homologs from human, mouse, rat, pig, and guinea pig using the PILEUP program [Wisconsin Package Version 9.1. Genetics Computer Group (GCG), Madison, WI]. The guinea pig sequences are partial sequences obtained by N-terminal sequencing of tryptic peptides (255). The signal sequences, EGF-like domains, proline-rich domain (mouse only), C1- and C2-like repeats, and putative amphipathic helix are indicated above the bovine sequence. Cleavage site for removal of signal sequence in bovine PAS 6/7 (↓), N- or O-linked glycosylation sites in bovine PAS 6 only (♦), O-linked glycosylation site in PAS 7 only (♦), N-linked glycosylation site in both PAS 6 and 7 (181) (♦), conserved C-residues (*), and RGD motif discussed in the text (v). GenBank™/EMBL or Swiss Protein (Sw. Prot.) Accession numbers are as follows: bovine, X91885; human, Q08431 (Sw. Prot.); mouse, M38337; rat, D84068; and pig, P79385 (Sw. Prot.).
prise two EGF-like domains between residues F251 and E6844 and between H6945 and T11389 (181) (Figure 9b). The RGD adhesive motif lies within the second of these two domains between residues R9167 and D9369 (arrowheads, Figure 9b). Each EGF domain is characterized by six C residues, which are conserved in the other characterized PAS 6/7 proteins (asterisks, Figure 9b) and in other EGF-like proteins, such as Drosophila Notch 1, a human proteoglycan core protein, and the neurogenic repetitive locus proteins 95F and delta (382). Because bovine PAS 6/7 contains no free sulphydryl groups (181), it can be assumed, by analogy with similar EGF proteins, that the C residues in each EGF domain form disulfide bonds. The predicted disulfide bonds in the first EGF domain are between C308247 and C295234 and C299238 (181). The C2-like domain, between residues Q211150 and C10278 and C11187 (181).

In bovine PAS 6/7, the EGF domains are immediately followed by tandem repeats between residues T15190 and N311250, and residues G312251 and the C-terminal, C470409 (181). These regions are, 60 to 63% similar to the C1 and C2 domains of the blood clotting factors V and VIII (181). Mass spectrometry of characterized peptides, established that the C1-like domains of bovine PAS 6/7 contain disulfide bonds between C15231 and C308247, and C295234 and C299238 (181). The C2-like domain contains a single disulfide between C313252 and C470409. A splice variant lacking 52 amino acids within the C1-like domain, between residues Q211150 and I264203, has been identified from a cloned cDNA (181); Figure 9a). This variant can be detected as a minor immunoreactive band in Western blots of bovine MFGM and is probably the consequence of exon skipping during mRNA synthesis (301).

The C2 domain of factors V and VIII (113, 299, 300) and the similar region of the PAS 6/7 proteins bind to phospholipids (7, 71, 181, 321). Binding is mediated via the C-terminal portion, which in each protein is predicted to form an amphipathic α-helix (127, 181) (within residues P454393 and L468407 for bovine PAS 6/7; Figure 9b). The hydrophobic face of the helix is presumed to insert into the interior of the phospholipid bilayer. However, in all characterized sequences, two charged residues, H458397 and R460399, interrupt the hydrophobic surface (Figure 9b). Andersen et al. (7) propose that any instability introduced by incorporation of the H residue into the bilayer could be stabilized by electrostatic interactions between the R residue and the phospholipid head groups. In current models, binding of the C-terminal domain of PAS 6/7 to the membrane bilayer, leaves the protein free to bind to integrin receptors on neighboring cells (or within the same membrane?) via the conserved RGD sequence in the second EGF domain [(7, 321, 394); Figure 9a, b).

Although all of the characterized, PAS 6/7 protein sequences predict the presence of both EGF and C1/C2 factor V/VIII-like domains, there are some significant differences in domain structure across species. Bovine PAS 6/7 is most similar to the putative pig and rat homologs in domain structure: two EGF-like domains fused to the tandem C1/C2 repeat (14, 110, 181, 295). Human PAS 6/7 lacks the first EGF-like domain but is otherwise similar to the bovine, pig and rat homologs [(91, 237), Figure 9b). From the limited peptide sequences available, the guinea pig protein (GP-55) appears to contain at least one EGF-like domain (similar to the N-terminal domain of bovine, mouse, pig, and rat PAS 6/7) and the tandemly repeated C1/C2 domains [(255), Figure 9b).

Two mouse homologs have been described, one with two EGF-like domains fused to the tandem C1/C2 repeat, like bovine PAS 6/7, and a second in which there is an additional P-rich region between the second EGF-like domain and the C1-like region [(110, 301, 382); Figure 9b]. These two forms arise from spliced variant mRNA in which the P-rich region, encoded by a single exon, is spliced out of one of the two mRNA (301). Interestingly, significant expression of the form containing the P-rich domain appears to be limited to lactating mammary tissue, suggesting tissue-specific functions (301). This domain contains six threonine residues and two serine residues that appear to be glycosylated with O-linked sugars (301). Currently, it is unclear whether the exon encoding the P-rich domain is present and expressed from the PAS 6/7 genes (86) of species other than the mouse.

Recent partial characterization of the mouse MFG-E8 gene suggests that the mosaic-like structure of PAS 6/7 proteins arose from exon shuffling during evolution and additionally, as indicated above, may be the consequence of exon skipping during mRNA synthesis. At least one EGF-like domain, the P-rich region, and the C1 and C2 domains are encoded by separate exons (301). The physiological basis for species differences in domain structure is not clear, although it should be emphasized that all of the currently characterized proteins contain the RGD adhesive motif in an EGF-like domain and a potential lipid-binding amphipathic α-helix in the C-terminus.

Posttranslational Modifications

Three independent groups have recently concluded that the two major bands of bovine PAS 6/7 [Band 15/16 (255), PAS 6/7 (181), MGP 57/53 (14, 15)] are differentially processed proteins with the same polypeptide
core. Kim et al. (224) also showed that two peptides purified from PAS 6 and PAS 7 samples were identical. However, in the same study, the HPLC profiles of tryptic peptides from each protein were substantially different. Despite this discrepancy, the weight of the evidence favors the possibility that bovine mammary PAS 6 and 7 are the products of a single gene, and that differences in size and isoelectric point are due to posttranslational modifications.

A major modification to PAS 6/7 is the addition of both N- and O-linked glycans. Removal of N-linked glycans with N-glycanase reduces both isoforms to a protein of approximately M_r 50,000 (15), close to the predicted M_r of 45,600 for the core polypeptide. Chemical analysis indicates that both isoforms also contain O-linked glycans (181). A number of isoelectric variants for each isoform have been described in several studies with pI ranging from 5.6 to 7.6 ([15, 224, 356]; Figure 2a, a'). Removal of sialic acid with neuraminidase reduces the number of variants to one for each isoform with approximate pI of 6.2 for PAS 6 and 6.5 for PAS 7 (224). Thus, the major source of heterogeneity (both size and charge) in PAS 6/7 proteins appears to be the addition of sialylated glycans.

Glycans associated with both PAS 6 (181, 223) and PAS 7 (181) have been characterized. Both isoforms contain N- and O-linked glycans (open and closed diamonds, Figure 9b) [however, see reference (223) for different conclusions]. In one study (181), the higher M_r isoform (PAS 6) was reported to be glycosylated on three residues: S339, N6541, and N270409, but the smaller, more abundant isoform (PAS 7) was only glycosylated on two residues, T4016 and N6541 ([181]; Figure 9a, b'). Compositional analysis and mass spectrometry of HPLC purified glycopeptides suggested that glycans bound to N270409 of PAS 6 comprise typical high-mannose structures. N6541 in both isoforms appeared to be glycosylated with hybrid-type glycans, most likely containing the unusual terminal GalNAc β1→4GlcNAc1 structure identified in bovine CD36 and BTN (see above). The O-glycans associated with S339 of the high M_r isoform (PAS 6) and T4016 of PAS 7 appeared to have similar compositions, comprising molar ratios of Gal, GlcNAc, and Fuc. A sialylated form was additionally identified in PAS 7.

In sharp contrast, no evidence was found for the unusual terminal GalNAc β1→4GlcNAc structure in PAS 6 in a separate study (223). Furthermore, the other characterized N-linked structures comprised complex-type glycans and a mannose-containing glycan with the structure Manα1→6(Manα1→3)Manβ1→4GlcNAcβ1→4GlcNAc. This latter glycan comprised 47% of the N-linked sugar recovered from PAS 6 and is the core glycan present in all N-linked glycoproteins, which in hybrid and complex structures is modified by addition of terminal sugars. Such high levels of the simple core glycan in a secreted protein is highly unusual and does not agree with the analysis of Hvarregaard et al. (181), who suggested that PAS 6 contains conventional high mannose-type structures. The reason for such markedly different results is uncertain.

To resolve these issues, more samples of PAS 6/7, collected at different stages of lactation, should be analyzed. Preliminary studies using lectins to determine the nature of the O- and N-linked glycans associated with PAS 6/7 and other MFGM proteins suggested that the glycosylation profile changes significantly during early lactation (15, 406). A reduction in sialylation was indicated in two such reports (15, 406) and N-acetylgalactosaminylation appeared to be stimulated (406).

Such developmentally dependent changes in sugar content clearly complicate comparisons between the glycosylation profiles of individual samples.

Other potential modifications to PAS 6/7, such as acylation and phosphorylation reactions, have not been reported.

**FABP**

FABP can be identified in the SDS-polyacrylamide gel profile of bovine MFGM as a protein band of approximately M_r 13,000 that stains with Coomassie blue and silver but not with the PAS reagent ([53]; Figure 1a, lane 1; Figure 1b, lane 1; Figure 1c, lane 5). A large fraction of the protein is recovered in the MFGM supernatant fraction ([53]; Figure 1a, lane 2; Figure 1b, lane 3; Figure 1c, lane 5). Gels containing over 10% (wt/vol) polyacrylamide are required for resolution of the protein, otherwise the protein migrates with other lower M_r material at the dye front (compare Figure 1a, lanes 1 and 2, with Figure 1c, lanes 5 and 6).

FABP was originally identified in screening assays for growth inhibitors of mammary carcinoma cells. Clarified supernatant fractions from lactating bovine mammary gland homogenates were shown to contain a potent inhibitor of the growth of Ehrlich ascites mammary carcinoma cells in vitro (47, 48). Inhibitory activity was purified using an in vitro cell proliferation assay to monitor biological activity and was identified as a protein of M_r 13,000, which was called mammary-derived growth inhibitor, (MDGI) (46, 47, 48). The entire sequence of this isolate was determined by Edman degradation (46). Comparison of this sequence with extant databases showed that MDGI was a member of the intracellular FABP family with closest similarity to rat heart FABP (46). Further work on the distribution of MDGI in bovine mammary tissue showed that the protein was present in soluble cell extracts and the MFGM
(both membrane and supernatant fractions) but only in trace amounts in milk whey (53). The growth inhibitory activity was subsequently shown to reside in a peptide of 11 amino acids derived from the extreme C-terminus (427). This peptide, the entire MDGI protein, or heart FABP were also shown to stimulate cellular differentiation of mammary cells, besides inhibiting cell division (427). Expression of MDGI mRNA and protein in vivo was positively correlated with glandular differentiation (427).

Doubts arose about the identity of MDGI because repeated attempts to clone a cDNA from mammary gland libraries were unsuccessful (49, 371). In fact, the cDNA that were initially isolated from mouse and bovine mammary gland libraries were identical to the cDNA that were initially isolated from mouse and bovine mammary gland libraries. However, the major FABP isolated from lactating rat or mouse mammary glands was shown independently by three groups to be heart FABP (29, 203, 286). A second FABP, the adipocyte form, was purified from virgin mouse mammary tissue (29). The abundance of this protein at early developmental stages was presumed to be due to the large amount of adipose tissue in prelactating glands. However, adipocyte FABP is also expressed during lactation, and a cDNA that is identical in sequence to adipocyte FABP was cloned from a lactating bovine mammary gland library (371). To complicate matters further, at least two additional FABP, keratinocyte lipid-binding protein (233) and an MDGI-related protein (355) that is identical to brain FABP (171), have been described in mouse and human mammary tissue, respectively.

An emerging consensus arising from these studies is that several FABP are expressed in mammary tissue, that in mouse the adipose form predominates at early stages of development, and that both adipose and heart FABP are expressed during pregnancy and lactation. During lactation in rat, mouse, and cow, heart FABP is most abundant (29, 49, 203, 286, 371). All of the FABP expressed in the mammary gland are potential “MDGI” because they are similar in the active C-terminal regions [reviewed in (171)]. In retrospect, the original MDGI isolates appear to have been a mixture of the heart and adipocyte proteins [(371); reviewed in reference (49)]. Perhaps most tellingly, the seven differences in amino acid residues between the sequence of MDGI and heart FABP are consenstant with the sequence of adipocyte FABP (Figure 10).

The MFGM form of FABP is most likely to be heart FABP because this is the most abundant tissue FABP during lactation (29, 203, 286, 371), and it is expressed in mammary secretory cells (236, 371), the presumed source of MFGM proteins. Furthermore, 27 of 29 residues identified in a limited number of peptides of “MDGI” isolated from bovine MFGM are identical to heart FABP [(53); Figure 10]. Nevertheless, it has not been rigorously shown that MFGM solely contains heart FABP, because only the tissue forms of mammary FABP have been extensively sequenced and characterized (29, 46, 203, 286), and all other sequences were obtained from cloned cDNA (43, 49, 233, 286, 371). Because of these considerations, the Committee recommends that use of the name MDGI be discontinued and that the generic term FABP be used to describe the MFGM isoform(s), until the exact identity of this protein in the MFGM is firmly established. It should also be noted that the general nomenclature of FABP is currently being revised. Descriptions based on tissue distribution are being replaced by the use of arabic numerals—FABP1 for “liver-type FABP” and FABP2 for “intestinal-type FABP,” etc. (49). This change is being promoted to avoid the confusion that has arisen from naming FABP after the first organ from which they were isolated, even though they are often more widely distributed in cells and tissues.

**Sequence Analysis**

FABP form a large family of over 20 proteins that have been sequenced and extensively characterized (27). The structures of at least 14 proteins have been determined by X-ray crystallography, including both
heart (431) and adipocyte (426) FABP. Detailed discussion of these structures is beyond the scope of this report, and the reader is referred to several reviews (27, 85, 346, 410), including an excellent discussion illustrated with stereo diagrams by Banaszak et al. (27).

FABP are folded into a barrel-like structure enclosing a central cavity formed by two sets of antiparallel $\beta$-strands that are twisted in orthogonal orientation with respect to each other. The $\beta$-strands are indicated by letters $\beta$-A, $\beta$-B, and so forth, in the sequences shown in Figure 10. Two subfamilies of proteins have been described. In the intracellular FABP, the barrel is compressed into a clam-shaped structure, and the cavity is capped by a helix-turn-helix motif (indicated by $\alpha$I and $\alpha$II in Figure 10). In the extracellular FABP or lipocalins, the barrel forms a nearly spherical structure; additional sequences are present at the N- and C-termini, and the helix-turn-helix motif is absent. Proteins in the intracellular subfamily typically comprise about 135 amino acids and include mammary FABP (heart and adipocyte proteins), liver and intestinal FABP, myelin P2 protein, and cellular retinoic-acid-binding proteins and retinol-binding proteins. Proteins in the extracellular lipocalin subfamily are composed of about 175 amino acids and include serum-retinol-binding protein, insect bilin-binding proteins, urinary $\alpha_{2}$-globulin, major urinary protein, and, most notably, the milk protein $\beta$-lactoglobulin (269). The overall structures within each subfamily are remarkably conserved, despite an overall low amino acid identity in many cases. Within the intracellular FABP family, internally disposed hydrophobic residues and adjacent water-exposed hydrophilic residues are conserved within the $\beta$-strands and within two of the tight turns that link contiguous antiparallel strands (27).

Rather unexpectedly, the cavity that accommodates the lipid ligands is not entirely hydrophobic in character. Ordered water molecules are present, and the residues lining the cavity comprise an approximately equal number of hydrophobic and polar amino acids. In most FABP, fatty acids and other lipid ligands bind within the cavity with the polar head groups oriented toward the base, farthest away from the opening. Key amino acids that are believed to function in ligand binding include conserved R or Q residues at positions 107$^{106}$ and 127$^{126}$ and a Y residue at position 129$^{128}$ (27) (asterisks, Figure 10). Gating of the cavity may be partially controlled by several residues including F58$^{57}$ (arrowhead, Figure 10), conserved in the heart and adipocyte isoforms, which adopt different conformations in the presence and absence of ligand (426).

The physiological functions of FABP are still not entirely clear (27, 346), despite extensive knowledge of the structures and ligand-binding specificities of many family members, including the mammary isoforms. The FABP in mammary tissue may function in the intracellular transport of fatty acids, control of lipid metabolism, or possibly the accretion of lipid droplets in the cytoplasm. Mammary FABP may also function in development by inhibiting the proliferation (47, 48) and stimulating the differentiation of epithelial cells (427) as originally observed by Grosse and colleagues. However, these experiments were conducted with ectopically added MDGI or heart FABP, which implies that the polypeptides function by signaling from outside of the cell (via receptors?). This possibility is difficult to reconcile with the intracellular location of the FABP, unless the proteins are exported across the plasma membrane by nonclassical (274) secretory pathways. The FABP secreted in association with the MFGM appears to be largely inaccessible to mammary cells because it is membrane bound, possibly via the cytoplasmically oriented termini of CD36 (376), or it is sandwiched between the lipid globule surface and the internal face of the membrane (53). Alternatively, mammary FABP may regulate growth and differentiation directly by intracellular mechanisms (49, 137, 171), and ectopically added “MDGI” could have been taken up in sufficient quantity to cause the developmental effects noted in cells grown in vitro.

Posttranslational Modifications

At least one minor and two major isoelectric variants of FABP in bovine MFGM can be resolved by two-dimensional gel electrophoresis (Figure 2c). As discussed above, it is most likely that these variants are different forms of bovine heart FABP, although in the absence of confirming data, the presence of minor quantities of adipocyte FABP or other FABP in MFGM cannot be discounted. Two variants of heart FABP, with pI of 4.9 and 5.1, are due to a single amino acid change at residue 99$^{98}$, which can be either a D or an N (371, 407), and these appear to be the most abundant variants in lactating bovine mammary tissue (371). However, the variants detected in bovine MFGM with a specific rabbit antibody to FABP from bovine mammary gland (376) have pI between 5.1 and 5.9 (Figure 2c and I. H. Mather, unpublished observations). Whether these differences are due to the use of, e.g., different electrophoresis systems or pI standards or whether they indicate different posttranslational processing reactions in individual animals or major differences between the tissue and MFGM forms for bovine FABP is not clear.

Bovine mammary FABP does not appear to be glycosylated. N-linked glycans are unlikely because the protein is synthesized in the cytoplasm (bovine heart FABP contains one putative NXS/T sequon). Neither N- nor O-
linked sugars have been detected in any of the purified intracellular FABP. Phosphorylation of residue Y20\(^{19}\) of heart FABP in rat heart myocytes and of adipocyte FABP in 3T3L1 cells by the insulin receptor gives rise to a minor acidic variant (67, 174, 287). A similar variant can be generated by stimulating freshly prepared rat mammary cells in vitro with insulin (286). Bovine FABP in MFGM is phosphorylated by intrinsic kinases when the membrane is incubated with \([\gamma^{32}\text{P}]-\text{ATP}\) (375). Therefore, at least some of the variants identified in Figure 2c may be due to differences in phosphorylation status.

Modification of mammary FABP by covalent addition of lipids has not been reported. Some “variants” may be due to the presence of residual lipid ligand, although under the denaturing conditions used to prepare samples for isoelectric focusing this is unlikely [see reference (346) for a comprehensive discussion of FABP isoforms and potential posttranslational modifications].

**Additional Protein Components**

Bovine MFGM preparations contain many more proteins and enzymes than those discussed above. These additional components include enzymes, immunoglobulins, MHC molecules, proteins derived from the cytoplasm of the secretory-epithelial cells, proteins from milk leukocytes, and skim milk constituents. A recent analysis of human MFGM by two-dimensional Western blot reveals how complex the protein composition may be (128). Many of these constituents are undoubtedly peripheral proteins loosely adsorbed to the membrane from the cell cytoplasm or from milk serum and are present in trace amounts. Whether these should be considered as membrane-associated proteins or contaminants often becomes a game in semantics. In many cases their exact cellular origin is uncertain. Constituents such as 5’-nucleotidase and phosphodiesterase are probably largely derived from the plasma membrane (317) and are presumably bona fide components of the MFGM bilayer. Other proteins such as acid phosphatase and catalase are normally considered to be components of lysosomes and peroxisomes, respectively. Such enzymes are elevated in milk from glands infused with endotoxin or from cows with mastitis (10, 11, 227, 229), and they are presumed to originate from polymorphonuclear leucocytes that have infiltrated into diseased areas of the gland from the systemic circulation (8). Leucocytes may ingest milk fat globules (335) and float with the cream layer during isolation of fat globules for preparation of the MFGM. Therefore, the increased content of acid phosphatase and catalase observed in the MFGM from mastitic milk is probably due to contamination from milk leukocytes. However, the source of such enzymes in the MFGM from disease-free glands is uncertain because leucocytes are rarely observed in cream from disease-free cows.

Mastitis may also lead to an increase in the content of enzymes in milk from the secretory epithelium because of physical damage. A large proportion of the enzyme N-acetyl-\(\beta\)-D glucosaminidase might be derived from secretory cells during udder infections, and a minor fraction might originate from leucocytes (227, 228). Again, this enzyme is present at low levels in the MFGM from disease-free cows and is of uncertain origin.

Some proteins may be adsorbed onto the outer membrane surface from the skim milk phase. Plasminogen bound to casein micelles apparently associates with bovine cream from skim milk and can be reduced to low quantities by washing fat globules at least three times with buffered salt solutions (32, 37). Some minor proteins of bovine MFGM may be derived from cytoplasmic material trapped between the outer membrane bilayer and the surface of the fat globule (“crescents”; 191, 421). However contamination of MFGM with cytoplasmic material is species dependent and is very low in dairy cows (191). The cytoplasmic enzyme, aldolase, may originate from such crescents (217).

Finally, some of the enzyme activities detected in bovine MFGM may not reflect the activities of single proteins but may be the consequence of multiple enzymes with overlapping substrate specificities (e.g., some of the hydrolases and redox components). Unfortunately, very few enzymes have been purified extensively from bovine MFGM; therefore, absolute amounts of specific constituents are generally unknown. A partial listing of additional protein components of bovine MFGM is given in Table 5. The Committee recommends that the standard nomenclature of the Enzyme Commission be adopted for the names of specific enzymes as they are identified and characterized.

**ACKNOWLEDGMENTS**

The author thanks members of the Milk Protein Nomenclature Committee for their advice and help in preparing this review (Committee members: R. Jimenez-Flores (Chairperson), California Polytechnic Institute, San Luis Obispo, CA; R. J. Brown, Utah State University, Logan, UT; H. M. Farrell, Jr., USDA, Philadelphia, PA; K. F. Kg Kwai Hang, Macdonald College, Quebec City, Canada; V. R. Harwalker, Center for Food and Animal Research, Ottawa, Canada; C. L. Hicks, University of Kentucky, Lexington, KY; D. J. McMahon, Utah State University, Logan, UT; and H. E. Swaisgood, North Carolina State University, Raleigh, NC).

An earlier version of this report was also reviewed by Kermit L. Carraway, University of Miami, FL; R.
Table 5. Additional proteins, enzymes, and other constituents associated with the bovine milk-fat globule membrane.1

<table>
<thead>
<tr>
<th>Protein</th>
<th>E.C. Number</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetyl-CoA carboxylase</td>
<td>6.4.1.2</td>
<td>361</td>
</tr>
<tr>
<td>N'-Acetyl-D-glucosaminidase (hexosaminidase)</td>
<td>3.2.1.52</td>
<td>8, 228</td>
</tr>
<tr>
<td>Acid phosphatase</td>
<td>3.1.3.2</td>
<td>102, 103, 218, 225, 226, 229, 263</td>
</tr>
<tr>
<td>Adenosine triphosphatase (activated by K+/Mg2+)</td>
<td>3.6.1.3</td>
<td>175, 218, 225, 317</td>
</tr>
<tr>
<td>Aldolase</td>
<td>4.1.2.13</td>
<td>103, 229</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>3.1.3.1</td>
<td>102, 103, 218, 225, 226, 229, 262, 263, 272, 273, 436</td>
</tr>
<tr>
<td>Angiogenin</td>
<td>...</td>
<td>168</td>
</tr>
<tr>
<td>Basic fibroblast-like growth factor</td>
<td>...</td>
<td>169</td>
</tr>
<tr>
<td>Catalase</td>
<td>1.11.1.6</td>
<td>229</td>
</tr>
<tr>
<td>Cholinesterase</td>
<td>3.1.1.7 (or 8?)</td>
<td>103</td>
</tr>
<tr>
<td>Cytochrome b5</td>
<td>...</td>
<td>23, 61, 192</td>
</tr>
<tr>
<td>Cytochrome P-420</td>
<td>...</td>
<td>61</td>
</tr>
<tr>
<td>Diaphorase (lipoamide dehydrogenase)</td>
<td>1.8.1.4</td>
<td>225, 273, 353</td>
</tr>
<tr>
<td>β-Galactosidase</td>
<td>3.2.1.23</td>
<td>11</td>
</tr>
<tr>
<td>Galactosyl transferase</td>
<td>3.2.1.22</td>
<td>326</td>
</tr>
<tr>
<td>Galactosyl transferase</td>
<td>2.4.1.38</td>
<td></td>
</tr>
<tr>
<td>Galactosyl transferase</td>
<td>2.4.1.90</td>
<td></td>
</tr>
<tr>
<td>Glucose-6-phosphatase</td>
<td>3.1.3.9</td>
<td>103, 225, 226</td>
</tr>
<tr>
<td>α-Glucosidase</td>
<td>3.2.1.21</td>
<td>8, 11</td>
</tr>
<tr>
<td>γ-Glutamyl transpeptidase</td>
<td>2.3.2.2</td>
<td>34, 225, 226, 262</td>
</tr>
<tr>
<td>Glycosaminoglycans</td>
<td>...</td>
<td>247, 358</td>
</tr>
<tr>
<td>GTP-binding proteins</td>
<td>...</td>
<td>126</td>
</tr>
<tr>
<td>UDP-glycosyl hydrolase</td>
<td>2.4.1.5</td>
<td>214</td>
</tr>
<tr>
<td>Inorganic pyrophosphatase</td>
<td>3.6.1.1</td>
<td>225</td>
</tr>
<tr>
<td>NADH/cytochrome c reductase</td>
<td>1.6.99.3</td>
<td>62, 103, 192, 218, 273</td>
</tr>
<tr>
<td>NADH-ferricyanide reductase</td>
<td>1.6.99.3</td>
<td>62</td>
</tr>
<tr>
<td>NADPH/cytochrome c reductase</td>
<td>1.6.99.1</td>
<td>62, 192, 218</td>
</tr>
<tr>
<td>5'-Nucleotidase</td>
<td>3.1.3.5</td>
<td>176, 218, 225, 226, 262, 317</td>
</tr>
<tr>
<td>Nucleotide pyrophosphatase</td>
<td>3.6.1.9</td>
<td>218, 225, 317</td>
</tr>
<tr>
<td>Phosphatidic acid phosphatase</td>
<td>3.1.3.4</td>
<td>64</td>
</tr>
<tr>
<td>Phosphodiesterase I</td>
<td>3.1.4.1</td>
<td>102, 103, 218, 262</td>
</tr>
<tr>
<td>Plasminogen/plasmin</td>
<td>3.4.21.7</td>
<td>37, 170</td>
</tr>
<tr>
<td>Protease peptone component-PP3</td>
<td>...</td>
<td>368, 369</td>
</tr>
<tr>
<td>Ribonuclease I</td>
<td>3.1.27.5</td>
<td>229</td>
</tr>
<tr>
<td>Sulphydryl oxidase</td>
<td>1.8.3.2</td>
<td>190, 225, 226</td>
</tr>
<tr>
<td>Thiamine pyrophosphatase (nucleoside diphosphatase)</td>
<td>3.6.1.6</td>
<td>218, 326</td>
</tr>
</tbody>
</table>

1Excluding milk serum proteins and caseins.
2Inactive enzyme, detected by immunoblotting.
3Protein purified substantially from bovine milk-fat globule membrane or colostral fat globule membrane.
4Traces.
5Protein purified from skim milk.

The author accepts all responsibility for sins of omission, mistakes, and bias. In a discussion of so many markedly different proteins, it was not possible to include all relevant literature, and, inevitably, some work had to be cited indirectly through reviews. The author apologizes to those not cited.

REFERENCES


121 Greenberg, A. S., J. J. Egan, S. A. Wek, N. B. Garty, E. J. Smart. 1999. The class B, type I scavenger receptor pro-


126 Ghosal, D., D. Ankrrap, and T. W. Keenan. 1993. Low molecular mass GTP-binding proteins are secreted from mammary ephe-


129 Graf, G. A., P. M. Connell, D. R. van der Westhuysen, and E. J. Smart. 1999. The class B, type I scavenger receptor pro-

130 Greenberg, A. S., J. J. Egan, S. A. Wek, N. B. Garty, E. J. Smart. 1999. The class B, type I scavenger receptor pro-


139 Harrison, R., and R. C. Bray, ed. 1997. Xanthine oxidase: enzy-


REVIEW: FAT-GLOBULE MEMBRANE PROTEINS


