

## BOVINE LACTOFERRIN mRNA: Sequence\*, Analysis, and Expression in the Mammary Gland

Richard E. Goodman<sup>+</sup> and Floyd L. Schanbacher

<sup>+</sup>J. A. Baker Institute, Cornell University, Ithaca, New York 14853

<sup>^</sup>Department of Dairy Science, the Ohio Agricultural Research and Development Center, Ohio State University, Wooster, Ohio 44691

Received August 10, 1991

**SUMMARY** The mRNA sequence for bovine lactoferrin expressed in the mammary gland was determined by sequencing three overlapping cDNA clones and by direct sequencing of the mRNA. The mRNA (2351 bases) codes for a 708 amino acid protein with a 19 amino acid signal peptide immediately preceding a sequence identical to the N-terminal 40 amino acids reported for bovine lactoferrin (1). A putative destabilizing sequence (AUUUA) was identified in the 3'-untranslated region. The nucleic acid sequence and deduced amino acid sequence are highly homologous with other transferrin family members. Lactoferrin mRNA concentrations in bovine mammary tissue were quite low two days before parturition and during lactation but were high three days after the cessation of milking, a sharp contrast from the pattern of regulation of the other milk proteins. © 1991 Academic Press, Inc.

Lactoferrin is a member of the transferrin gene family of non-heme iron-binding glycoproteins and is found in polymorphonuclear leukocytes and in mucosal and epithelial secretions (2). It is the major iron-binding protein of bovine and human milk and mammary secretions (3). Lactoferrin is recognized for its importance in the transport of iron to suckling young, as well as to the antimicrobial defenses of the maternal mammary gland and neonatal gut (3,4,5). It also modulates immune functions, granulocyte and lymphocyte mitogenesis and localization (6), and growth of intestinal epithelium (7).

Lactoferrin's regulation is contrary to that of other milk proteins. In mammary secretions, bovine lactoferrin (bLf) is elevated in colostrum (1-2 mg/ml), is lowest in milk (0.01-0.1 mg/ml), and increases remarkably (20-100mg/ml) during mammary involution as the secretion of milk and other milk proteins decrease (8). After intramammary infection, lactoferrin is elevated in milk (>2 mg/ml) for prolonged periods even though other milk proteins remain at normal concentrations (9). In human milk, lactoferrin concentra-

\* The sequence of bovine mammary lactoferrin mRNA has been accepted for deposit in the GenBank data bank under accession number M63502.

<sup>^</sup>To whom correspondence should be addressed at Lab. Mol. and Dev. Biology, Dept. Dairy Science, Ohio Agricultural Research and Development Center, Ohio State Univ., 1680 Madison Ave., Wooster, OH 44691.

**Abbreviations:** bLf, bovine lactoferrin; hLf, human lactoferrin; hTf, human transferrin; mLf, mouse lactoferrin; oTf, ovotransferrin; pTf, porcine transferrin; aa, amino acid.

tions are high in colostrum (1-40 mg/ml) and at 90 days of lactation remain higher (0.04-6.0 mg/ml) than in bovine milk (10).

For the transferrin gene family of iron-binding glycoproteins, cDNAs described for human serum transferrin (11), ovotransferrin (conalbumin) of egg white (12), and for human (13,14), and mouse (15, incomplete) mammary lactoferrins indicate they code for single polypeptide chains of about 700 amino acids, with an iron-binding site in each of the similar N- and C-terminal halves. In spite of high homology and structural similarities, little is known about the tissue or species specificity for regulation of lactoferrin expression (2). The structural basis for human lactoferrin's iron binding has been defined by X-ray crystallography (16), but the structure recognized by lactoferrin receptors has been localized only to the 282 amino acid N-terminal fragment (17). Recently described cDNA's for bovine lactoferrin, published after the completion of our work, are incomplete at the 5' end (18,19). We describe here the first complete sequence for bovine mammary lactoferrin mRNA (20) and show that its regulation is opposite that for other milk proteins during bovine mammary development and involution.

#### MATERIALS AND METHODS

Bovine mammary tissues were: developing gland (estimated to be two days prepartum, #3091); peak (117 days) lactation (#3017); late (318 days) lactation (#3148) with continued regular milking of one-half of the gland while the contralateral half was unmilked for three days to induce early involution; and early involuting mammary tissue, induced in the entire gland (70 days) by complete cessation of milking three days prior to slaughter (#2949). Cellular RNA isolated from cultured bovine mammary cells (21) and mammary tissue #2949 using the guanidinium isothiocyanate/CsCl method, was further purified with oligo-dT cellulose (22) and used in the construction of two cDNA expression libraries in lambda gt11 (Amersham and Invitrogen kits, respectively). The libraries were screened with rabbit antisera to bovine lactoferrin (21), detected with alkaline phosphatase labeled secondary antibody, and with <sup>32</sup>P-labeled cDNA probes derived from the first antibody positive clones. Inserts from the three lambda clones were subcloned into pGEM-4Z (p1f3a, p1f3ac, pa3, pb4, pc3) or pGEM-5Zf(+) (pn16b, p5'-5), then further subcloned by restriction and religation (Fig. 3).

Sequenase version 2.0 (U.S. Biochemicals) kits were used to sequence the cDNA clones by Sanger dideoxy-chain termination of denatured plasmids as described in the manufacturer's instructions, using T7 or SP6 primers. The 5' end of the mRNA sequence was determined by direct dideoxy-sequencing (23) of MeHgOH denatured mRNA, primed with one of three oligonucleotides (#1, 5'-AGCATCCGCCTTTTCTC-3'; #2, 5'-TCCCATAGATCTCTGCTG-3'; and #3, 5'-GTCCCGGCCCGCCTCAAAC-3'), and sequenced using a GemSeq, AMV reverse transcriptase sequencing kit (Promega). The [ $\alpha$ -<sup>35</sup>S]dATP labeled sequencing products were detected following electrophoresis in 6% polyacrylamide, 7M urea sequencing gels (22), by autoradiography with XAR X-ray film (Kodak). Homology comparisons and protein structural analysis were performed with ALIGN and PROTYLZE software, respectively, from Scientific and Educational Software (State Line, PA).

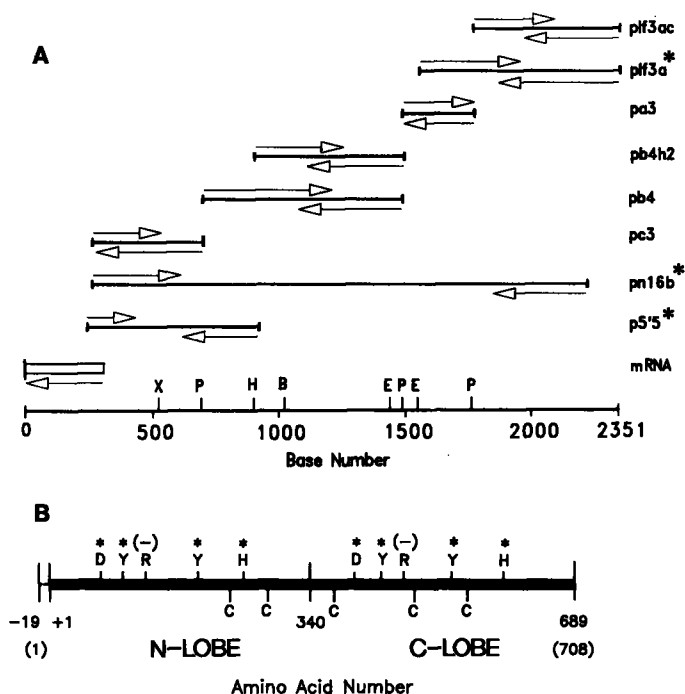
Northern blots were performed with cellular RNA from bovine mammary tissues (above), liver and circulating leukocytes (buffy coat of venous blood); isolated by the guanidinium/phenol method with 4M LiCl precipitation (24). RNA (20 ug/5 ul H<sub>2</sub>O) was diluted with 25 ul of denaturing solution (mM Na HEPES, pH 7, 6 mM NaOAc, 1.2 mM NaEDTA, 60% formamide [deionized], and 7.2% formaldehyde [deionized]), heated to 70°C for 10 mins. and electrophoresed through 1.5% agarose (0.82 M formaldehyde), with 20 mM sodium phosphate (pH7.0) running buffer. The gel was then stained with ethidium bromide before the RNA was denatured in NaOH, neutralized with Tris and transferred by capillary blotting with 10X SSC (22) to GeneScreen Plus, nylon membranes. Membranes were baked at 80°C for two hours, then prehybridized with BLOTTO (25) and hybridized with 200 ng of random primed, [ $\alpha$ -<sup>32</sup>P]dCTP labeled cDNA

insert (22). Bovine cDNA probes were: pLf3a (bLf);  $\alpha_{s1}$ -casein (26); beta-casein, pC14 (27); kappa-casein, pC371 (27); beta-lactoglobulin, pBL13 (28); and alpha-lactalbumin, pALac (K. Brew, Miami Univ., unpublished). Post-hybridization washes were 2X SSC, 0.1% SDS at room temperature, followed by 0.2X SSC and 0.1% SDS at 55°C before exposing XRP X-ray film with the membrane using intensifying screens at -70°C.

## RESULTS AND DISCUSSION

The complete sequence of bovine lactoferrin mRNA isolated from mammary glands was determined by sequencing three over-lapping cDNA clones and sub-clones (bases 237-2351) and by direct sequencing of the 5'-terminus of the mRNA (bases 1-237) as indicated in Figure 1. Clone pLf3a was isolated by screening  $3 \times 10^5$  plaques from the library derived from cultured mammary cells, using anti-bLf serum. An open reading frame in the 5' end of pLf3a was determined to have 70% identity with the 3'-coding regions of mouse lactoferrin (15) and human transferrin (11). Rescreening ( $10^5$  plaques) by hybridization with  $^{32}$ P-labeled cDNA from pLf3a failed to identify any longer clones. The second library (unamplified) was screened with anti-bLF serum ( $2.5 \times 10^5$  plaques) and by hybridization with the labeled insert of pLf3a ( $1 \times 10^5$  plaques). The sequence of the 3'-end of the longest of 44 clones, pN16b, matched the 5'-end of pLf3a. The second library was rescreened ( $5 \times 10^5$  plaques) with a labeled 5'-Xho I fragment of pN16b, yielding just one clone (p5'-5) which extended further 5' than pN16b. The combined cDNA sequence was incomplete based on the estimated protein length for bLf. The 5'-end of the sequence was completed by dideoxychain termination sequencing of RNA template using oligonucleotide primers chosen from segments of the sequenced cDNA. We used MeHgOH to denature the RNA template to prevent premature termination of the reaction. Even with this procedure the reaction with the most 3' of the three primers was unsuccessful, probably due to strong secondary structures which could block the AMV reverse transcriptase. Analysis of the mRNA sequence using the RNAFold program predicted a complex secondary structure between bases 5 and 305 with a bond energy of -81.6 kcal/mol. The sequencing method used here for mRNA sequencing is analogous to primer extension (29) and consistently terminated in an unreadable series of 4 bases indicating a probable cap structure. Three bases in the sequence were ambiguous but were assigned ( $C_{213}$ ,  $C_{217}$ , and  $U_{235}$ ; Figure 2) on the basis of sequence data published for other lactoferrins. The indicated bases  $C_{213}$  and  $C_{217}$  are the only ones which would code for arginine (aa 39) and alanine (aa 40), respectively, identified at those positions by previous sequencing of the 40 amino acid N-terminus of mature bovine lactoferrin (1). The assignment of base 235 as U was used since Ile (aa 46) is conserved at this position in mouse (15) and human lactoferrin (30), human transferrin (11), porcine transferrin (31), and hen ovotransferrin (12).

The bLf mRNA (Fig. 2) includes a 42 bp 5'-UTR, a continuous open reading frame coding for a 708 amino acid signal peptide, and 185 bases of 3'-UTR containing putative destabilization signals (32,33) of AUUUA and AUUUUA at 32 and 12 bases, respectively, 5'-of the polyadenylation signal at base 2325, and a polyadenylation tract starting at base 2347. The sequence around the translation initiation codon at base 42, AGCCCAUGA, with a C-3 bases from the AUG is unusual for its occurrence in only 2% of the 699 mRNA's examined by Kozak



**Figure 1.** Diagram of cDNA plasmid clones, cDNA and mRNA sequencing strategy, restriction enzyme sites and deduced protein features for bovine mammary lactoferrin.

**A.** Direction and extent (arrows) of sequence determinations for each cDNA clone (bold lines), with the three independent primary cDNA clones (\*) shown below the secondary cDNA clones derived from each. The 5'-terminus was determined by direct mRNA sequencing (open box). Restriction endonuclease cleavage sites are: B, Bam HI; E, Eco RI; H, Hind III; P, Pst I; and X, Xho I.

**B.** Protein features indicated for the 708 amino acid sequence deduced for the complete lactoferrin protein include: signal sequence (-19 to -1), amino acid residues (single letter code) involved in iron (\*) and anion (-) binding; potential glycosylation sites (c); and the expected N- and C-lobe division near amino acid 340.

(34) and differs substantially from the optimum initiation sequence of CCA/GCCAUGG. Comparison of our bLf sequence with the recently published incomplete bLf sequences (18,19) identified a small number of mismatches (Fig.2). The additional 13 bases between the poly(A) signal and the start of the poly(A) tract which appears in the other bLf sequences may be due to alternate splicing. A consensus of the three sequences (not shown) was determined for bLf and used to deduce an amino acid sequence (Fig. 3) for comparison with other members of the lactoferrin/transferrin family. This consensus polypeptide differs from the deduced sequence of our bLf at amino acids (consensus/deduced) 47 (R/P), 48 (A/G), and 278 (F/S). Whether sequence differences represent genetic diversity or experimental error isn't known.

Our nucleic acid sequence coding for the mature protein and the deduced consensus amino acid sequence of the mature protein of bLf are highly homologous with published sequences for human lactoferrin (77% and 68%, respectively) and to lesser degrees with mouse lactoferrin (72% and 64%), human transferrin (68% and 60%), porcine transferrin (67% and 61%), ovotransferrin (62% and

	Met	signal pep >	
bLf	NNNNGAGCCUUCGUCCGGAGUCGCCCCAGGACGCCCAUGAGCUUCGUCCCCGCCUCCUGGAGCCUUGGACUGUGUCUGGCGUC		( 100)
P-S	.....G.....		( 50)
bLf	CCCGAGGAAAAACGUUCGAUGGUGUACCAUCUCCCAACUGAGUGGUUCAAUAGCCGAGAUCCGAGUGGAGGAUGAAGAAGCUGGGGUCUCCUUAUC		( 200)
P-S	.....C.....C.....		( 150)
M+T	.....C.CC.....CGC.....		( 80)
bLf	ACCUGUGAGAGGCGGCCUUUGCCUUGGAAUGUUAUCCGGCAUCGCGAGAAAAAGCGGAUGUGUACCCUGGAUGGUGCAUGGUGUUGAGGCGG		( 300)
P-S	.....A.....C.G..C.....		( 250)
M+T	.....A..AA.....C.G..C.....		( 180)
bLf	GCCGGGACCCUJACAAACUGCGGCCAGUAGCAGCAGAGAUUAUUGGACGAAAGAGUCCUCCCAACCCACUUAUUGUGUGGCCUGUGUAGAAAGGG		( 400)
bLf	CAGCAACUUCAGCUGGACACGUGCAAGGCCGGAAGUCCUGCCAUACGGGCUUGGCAAGUCCGUGGGUGGAUCAUCCUUAUGGGAUCCUUCGCGCG		( 500)
P-S	.....G.....		( 450)
bLf	UACUUGAGCUGGACAGAGUACUAGCAGCCUCCAGGAGCUGUGGCUAAUUCUUCUUGCCAGCUGUGUCCUUGCAUUGAUAGACAAGCAUACCCCA		( 600)
bLf	ACCUGUGUAAACUGUGCAAGGGGAGGGGAGAACAGUGUGCCUGUCCUCCCGGGAACCAUACUUGGGUUAUUGUGGCCUUAAGUGUCUGCAGGA		( 700)
bLf	CGGGGUGGAGAGCUGGGCUUUGUUAAGAGACGACAGUUGUAGAAUCCAGAGAAAGGUGAGCAGGAGCCAGUAGAGCUUCUUGCCUGAACAAAC		( 800)
bLf	AGUGGGGCGCAGUGGAUGCCUUAAGAGUGCCACUUGGCCAGGUCCUUCUUAUGCUGUGGCCGGAAGUGUGGAUGGCAAGGAAGACUUGAUUCU		( 900)
bLf	GGAAGCUUCAGCAGGAGGCGAGGAGAAUUGGAAAAACAAGUCCUGGAGCUUCCAGCUCUUGGCUUCCACCCGGCCAGAGGGACCUUGCUUCAA		(1000)
P-S	.....U.....		( 950)
M+T	.....U.....		( 880)
bLf	AGACUUGUCUUGGGUUUUGAGGAUCCUUGCAGGUAAGUUGGCGUGUACUUGGCCUCCCGCUUUGACCACCUUGAAGAACCUAGGGAACU		(1100)
P-S	.....U.CG.....		(1050)
bLf	GCGGAGGAGUGAAGGCGCGUACACAGGUCUGUGUGUGCCUGGGACCUAGAGGAGCAGAAGAAGUGCCAGCAGUGGAGCCAGCAGAGCGGCGAGA		(1200)
P-S	.....C.....		(1150)
bLf	ACGUGACCUUGGCCAGGCGUCCACCACUGACGACUGCAUCUGUCCUGGUGUGAAAGGGGAAGCAGAUCCCUGAACUUGGAUGGAGGAUUAUUAUACAC		(1300)
P-S	.....C.....		(1250)
bLf	UGCGGGCAAGUGUGGCCUGUGGCCUGUCCUGGCAAGAACCGGAAUCCUCAAACACAGUAGCCUAGAUUGUGUGCUGAGACCAACGGAAGGGUACCUU		(1400)
P-S	.....U.....A.....		(1350)
bLf	GCCUGGCGAUUGUCAAGAAAGCAAUAGGGGCUACAUUGAAUUCUUGAAAGACAAGUCCGACACCCGGCGUGGACAGGACUGCAGGCGUGGA		(1500)
bLf	ACAUCCCCAUGGGCCUGAUCUGCAACACAGACAGGCCUCCUGCGCAUUGAUGAAUUCUUAUGACAGAGCUGUGCCCCUGGGGUGACCCGAAUCCAGACU		(1600)
P-S	.....CG.....		(1550)
bLf	CUGUGCCUUGUGUGUGGCGAUGACAGGGGCCUGGACAAGUGUGGCCAACUCUUAAGGAGAAGUACUAGGCUUAUACCGGGGCUUACAGGUGCCUGGU		(1700)
bLf	GAGGACGUUGGGGACGUUGCCUUGUGUAAAAACGACACAGUCUGGGAGAACACGAUUGGAGAGAGCACUGCAGACUGGGCUAAGAACUUGAAUCUGGAGG		(1800)
bLf	ACUUCAGGUUGUCUGGCCUGAUGGCACAGGAAGCCUGUGACGGAGGUCAGAGUGGCCACUUGCGGGUGGCCCGAAUACGCUUGGUGUGUCUGGAG		(1900)
P-S	.....C.....		(1850)
bLf	CGAUAGGGCAGCACACGUGAAACAGGUGCGUCCACCAGCAGGCUUGUUGGGAAAAUUGAAAAUCCGCCGACAAUUGUUGUUGUCAAUUAUCU		(2000)
bLf	GAAACCAAAAACCUUCUGUCAAUGACAACACUGAGUGUGGCGAAACUUGGAGGCAGACCAACGUAUGAAGAAUUAUUGGGGACAGAGUAGUCACGG		(2100)
	END		
bLf	CCAUGGCCAACUGAAAAAUGCUAACUCCCCGCUUCUGGAAGCCUGCGCCUCCUGACGAGGUAAAGCCUGCAAAAGAGCUAGCCUGGCCUCCUGGG		(2200)
bLf	CCUCAGCUCCUCCUGCUCUAGCCCCAAUCCAGGCGGAGGACCUUCCUCCUCCUUGGAAGUGCGGAUUUUUGCCAAGCUACAGUAUUUACAA		(2300)
	=====		
	+++++		
bLf	UUCUCUGCUGUUAUUUAGCAAGAAUAAAUUAGAAUUGCUUG-----AAAAA		(2351)
P-S	.....AUUUCAUCCCU.....		(2314)
M+T	.....C.....AUUUCAUCCCU.....		(2244)

**Figure 2.** Nucleotide sequence of bovine mammary lactoferrin mRNA and comparison with newly published bovine lactoferrin cDNA's. Sequences of the complete bLf mRNA (20) and incomplete cDNA's, P-S (18) and M+T (19) are listed (differences only). The cDNA's are converted to mRNA for comparative purposes. Gaps (-) and identical bases (.) are marked. The initiation codon (Met), the end of the signal peptide (>|), the termination codon (END), putative destabilization signals (\) polyadenylation signal (=) and poly A tract (+++) are identified. Undetermined nucleotides are indicated (N).

53%), and melanotransferrin (35, not shown) (57% and 41%). The 19 amino acid signal peptide sequence coded from the initiation codon at base 42 conforms to the common patterns of signal peptide sequences (36) and should be cleaved at

-19	-1	+5	s	s	s	s	s	*
blf	MKL	FVPALLSLGALGLCLAAPRKNV=RWCTISQPEWFKRRWGRMKKLG=	PSITCVRRFALEICIRAI	AEKKADAVTL	DGGMVFEAGRDPYKLRPV	( 77)		
mLf	==.LI.S.IF.E.....=	KATT.=...AV.NS.EE..L...NE.R.V.G==	PLS..KKSSTRQ.Q..VTNR..M.....TM.D..KP.....	( 77)				
hLf	...VFLV.F.....GR.RS.=Q..AV.N..AT..FQ..RN.R.VRG==	PVS.LK.DSPIQ.Q...NR.....FIY...LA.....	( 77)					
hTf	.R.A.G...VCAV.....V.D.T.=	AV.EN.AT..QSFRDH..SVIPSDG..VA..KK.SY.D.....AME.....A.L.YD.YLA.NN.K..	( 80)					
pTf	=====VAQ.T.=	NQ.AN..SSFREN.S.AVKNG=LVS..KKSSY.D..K..RD.E.....A.L.....LA..N.K..	( 79)					
oTf	...ILCTV...IAAV.F...P.S.I.....S..EK.=NNLRDLTQER==	I.L..QK.TY.D..K...NNE...IS...Q.....LA...K.I	( 77)					
blf	AAE	IYGTGESQTHYYAVAVVKGSFQDLQGRKSCHTGLGRSAGWIIPMGILRPYLSWT=====	ESLEPLOGAVAKFFSASCVPCIDRQAYPNLC	(170)				
mLf	...V...Q.R.....NS...H.N...LR.....I.....K..I.T.....N.N=====	GPPAS.EE..S...K...GAQKDRF....	(170)					
hLf	...V...ERQ.R.....GS...NE..L.....R.T...NV.I.T...F.N.....	GPP..IEA...GA.KGQF....	(170)					
hTf	V..F..S.D..F.....D.G..MN..R.K.....N..I.L.=	CDLP=====PRK..EK...N..G..A..A.GTDF.Q..	(171)					
pTf	V..F..Q.DN.....WN..KR.....L.=	DQLP=====PRK.IEK...S...A.PVNF.K..	(170)					
oTf	...EHT.GST.S.....TE.TVND...KT.....N..I.T.=	H.GAIEWEGI..GSVE.=	==GATIEQK..	(171)				
blf	QLCKGEGENQACSSREPYFGYSGAFKCLQDGAGDVA	FVFKETTVFENLPEKARDQYELLCLNNSRAPVDAFECHELAQVPSHAVVARSVDGKEDLIWKL	(270)					
mLf	SS.A.T.A.K.S.PE..S.A..LR..R.N.....TRGS..E.N.E..K...PD.TWK.TEY.....S.TND..EA.E.	(270)						
hLf	R..A.T..K..F..Q...S.....IR.S..D.SDE.E..E...PD.T.K..K..D..R.....N...A.N.	(270)						
hTf	..P.=====G..TLNQ.....K.....HS.I..AN.....D.T.K..EY.D.....T...MG.....E.	(266)						
pTf	.Q.A.K.AEK..NH.....A..N..KED.....HS.L..D.....RD.T.R..DYEN.Y.....Q..S.E.	(270)						
oTf	RQ...DPKTK..=RNA..S.....H..K.K.K.....H..N.A.=DLN.E....DG..Q..NY.T.NW.R.AA.....D.N.VED..SF	(266)						
<< N-LOBE C-LOBE >>								
blf	LSKAEKFGKNSRSFQLFGSP=====PGQRDLFKDSALGFLRIPSKVDSALYLGSRYLTLTKNLRETA==	EEV==KARYTRVVWCAVGPPEEKQKQWS	(362)					
mLf	.RQS.....KQASG..A.=====S..K.....E..I..V.V.Q..VG...TFS.T.SIQ..NKKQ==QD==	I.SKA..T.....S..KR..D..N	(362)					
hLf	.RQ.....D..PK.....=S..K.....I..S.V.PRI..G...G.F.AIQ..KSE==	A..RA.....EQ.LR..N..	(362)					
hTf	.NQ..H..D..KE..S=====H..K.....H...KV.PR.M.AKM..YE.V.AIR...GTCP.AP=TDECKP.K..LSHH.RL..DE..	(359)						
pTf	.NQ..H..RD..PD...S.S=====H..K.....N..K..M..S...YQ.V.A.R...EI=SPDSS.NECKK.R..I.H..TQ..DA..	(363)						
oTf	...SD..VDTKSD.H..P.GKKD.VLK.....IMLK.V..LM..Q...FE.YSAIGSM.KDQ==LTP=SP.EN.IQ...KD.KS..DR..	(362)						
blf	QQSGQNVTCATASTTDCIVLVKGEADALNDGGYIYTAGKCGLPVPLAENRKSSKSSLD=CVL	RPTEGYLAVAVVKKANEG=LTNLSLKDKKSCHT	(459)					
mLf	RD.RGR...ISFP..E..AIM..D..MS.....Q...SNG...=N..V...A.RREDA=F..S..RG.	(459)						
hLf	GL.EGS..SS..E..A..MS...V.....Y..QQS.DP.PN..D..V.....RRSDTS=..V.G.	(461)						
hTf	VN.VGKIE.VS.E..E..AKIMN..MS...FV.I.....=YNK.DN==EDT.EA..F.....SASD=..DN.G.	(452)						
pTf	IN..GKIE.VS.EN.E..AKIV..MS...I.....=Y.TEEN==NT.EK.....SSGP.D..N..G.	(458)						
oTf	VV.NGD.E.TVDE.K...=KIM...VA...LV...V...M...=RYDDE.QCS=KTDERPAS.F...AR.=DSN=VN..N..G.	(455)						
blf	AVDRTAGWNIPMGLIVNQTGSCAFDEFFSQSCAPGADPKSRLCALCAGDD=QGLDK=CVPNSKEKYGYTGAFRC	LAEDVGDVAFVFKNDTVWENTNGEST	(557)					
mLf	.....LA..R..K.N.....N.....I..E=K..EN=.A...R.Q...L...KA.N...L.DS..LQ..D.KN.	(557)						
hLf	.....LF.....K..Y.....S..R.N...I..E..EN=.N..R.....N.R.....DV..LQ..D.NNN	(559)						
hTf	.G.....LY.KINH.R...EG...SKKD.S..K..M...=S.NL=E..N..G.....V=K...HQ..PQ.TG.KNP	(547)						
pTf	.....LY.KIN..K..Q..EG...SORN.S...I..SE=RAPGRE.LA.NH.R.....V=K...DOV.QQ..D.KNK	(556)						
oTf	.G...V.....H.R..T.N..Y..EG...SP.N...Q..Q.SGGIPPE.=AS.H...F...L..V=K...IQHS..E..G.KNK	(553)						
blf	ADWAKNLNREDFRLLCLDGRKPVTEAQSCHLAVAPNHAVVSRSRAAHVKQVLLHQQALFGKNGKNCPOKFC	LFKSETKNLLFNDNTECLAKLGGRPY	(657)					
mLf	EE..R..KLK..E...D...KN.....T.KVEVLQ..V.D..VQ..R..QR..GE...Q.K.....IP.KT.S	(657)						
hLf	EA..D..KLA..A...K...R...M.....M.KVERL.....K..R..SD.....R.H.KT..	(659)						
hTf	DP...EK.YE...E..YAN...R...T.K.KE.C.HKI.RQ..H...S.VTD.SGN...R...D..R.D.V...RD.N..	(647)						
pTf	D...D.KQM..E..QN.A.E..DN.EN...R...A.D.KVTC.AEE..K..Q..RHVTD.SSS..M..N..D..R.D.Q...HV.KT..	(655)						
oTf	.....QMD..E...T..R.AN.MDYRE.N..EV.T...V.PEK.NKIRDLER.EKR..V...=SEKS..MM.E.QNKD...K.L.K..F.VREGT..	(652)						
blf	EEYLGTEYVTAIANLKKCSTSPLEACAFL=====	T=R (689)						
mLf	.K...K...I.TER..Q..S.....	=Q (689)						
hLf	.K...PQ..AG.T.....E.....	=R=K (691)						
hTf	.K...E...K.VG..R...S...T.R=====	R=P (679)						
pTf	.S...AD..I..V...R...K...T.HSAKNPRVET.=	T (696)						
oTf	K.F..DKFY.V.S...T.NP.DI.GM.S.....	=EGK (685)						
				Homology to blf				
				---				
				64%				
				68%				
				60%				
				61%				
				53%				

**Figure 3.** Amino Acid Homology Alignment of transferrins/Lactoferrins. Actual (hLf) and deduced (bLf, mLf, hTf, pTf, and oTf) amino acid sequences indicate conservation of the spatial arrangement of the iron binding (\*) and anion binding sites (-). Cysteines probably involved in disulfide bridges (s) are highly conserved. Potential glycosylation sites for bLf (c); identical amino acids (.) ; gaps (=) and the division between lobes (N-LOBE \ C-LOBE) are indicated. Amino acid numbers for signal peptides are negative while those for the mature protein are in parentheses. Abbreviations and references are shown: bLf, bovine lactoferrin (consensus of 18,19,20); mLf, mouse lactoferrin (15); hLf, human lactoferrin (13,14); hTf, human transferrin (11); pTf, porcine transferrin (31); oTf, ovotransferrin (12). The consensus bLf sequence differs from our sequence at three amino acids (θ).

**Table 1.** Comparison of predicted disulfide bonding for bovine lactoferrin and human lactoferrin and transferrin. Intra-peptide disulfide bridges, identified as probable for human lactoferrin (30), with modified amino acid numbers (16) to match the cDNA corrected primary sequence. Cysteines found in identical positions in the deduced amino acid sequence of bovine lactoferrin are predicted to participate in the corresponding disulfide bond, with bond numbering as adopted from Metz-Boutigue et al. (30) preceded by N for those of the N-Lobe and by C for those of the C-Lobe.

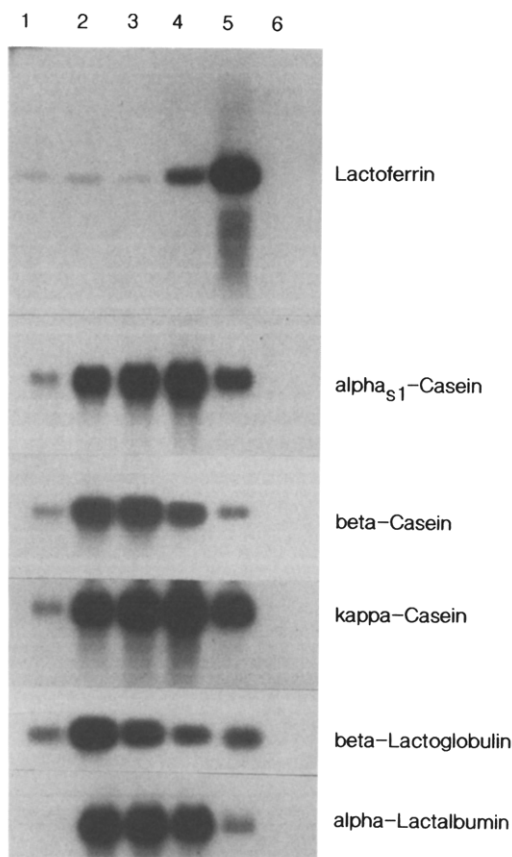
	Disulfide Bond No.	Human Lactoferrin Residue No.	Bovine Lactoferrin Residue No.	Human Transferrin Residue No.
N-Lobe	N1	9 - 45	9 - 45	9 - 48
	N2	19 - 36	19 - 36	19 - 39
	N3	115 - 198	115 - 198	118 - 194
	N4	157 - 173	157 - 173	158 - 174
	N5	170 - 181	170 - 181	171 - 179
	N6	231 - 245	231 - 245	227 - 241
	N10	-----	-----	137 - 331
	N11	-----	160 - 183	161 - 177
C-Lobe	C1	348 - 380	348 - 380	345 - 377
	C2	358 - 371	358 - 371	355 - 368
	C3	459 - 534	457 - 532	450 - 523
	C4	493 - 507	491 - 505	484 - 498
	C5	504 - 517	502 - 515	495 - 506
	C6	575 - 589	573 - 587	563 - 577
	C7	483 - 677	481 - 675	474 - 665
	C8	427 - 649	425 - 647	418 - 637
	C9	405 - 686	405 - 684	402 - 674
	C12	-----	-----	339 - 596
	C13	627 - 632	625 - 630	615 - 620

the Ala<sub>19</sub>-Ala<sub>20</sub> juncture during translation to yield the mature protein of bLf with an N-terminus identical to that reported by Rejman (1).

Amino acid residues involved in iron and anion binding in human lactoferrin by Anderson (16) are conserved in each of the two halves of bLf, as in all other known lactoferrins and transferrins (Fig. 3). Other amino acids immediately surrounding arginines 121 and 463 (ThrXXXArgXAlaGly) which may also participate in anion and therefore iron binding, as predicted for hLf (16), are also conserved in bLf, hLf, mLf, hTf, pTf, and oTf. The bLf sequence presented here includes five potential N-linked glycosylation sites; Asparagines 233, 281, 368, 478, and 547. Bovine and Caprine lactoferrins are thought to have four glycan chains normally (37).

Cysteine residues and disulfide bonds established for ovotransferrin and human transferrin and lactoferrin (30), are highly conserved in this family of proteins (Fig. 3). Based on the deduced amino acid sequence, bLf likely has a total of 17 disulfide bonds (Table 1), including the 16 disulfide bonds present in hLf plus one additional in the N-lobe of bLf which is found in hTf (N11) but not in hLf. Two additional bridges (N10 and C12) of hTf are apparently absent from both bovine and human lactoferrin.

Northern blots of RNA from liver (negative control, not shown), peripheral circulating leukocytes, and mammary glands during prepartum development, lactation, and early involution (Fig.4) showed the expected sizes (estimated from 28s and 18s rRNA) for each of the milk protein mRNAs shown, and indicate



**Figure 4.** Northern blot analysis of mammary and leukocyte mRNA for lactoferrin and milk protein expression. Cellular RNA (20 ug/lane) was from bovine mammary tissues [two days prepartum, #1; peak lactation, #2; late lactation, milked half, #3; involuting late lactation, unmilking for 3 days (same cow as #3), #4; involuting mid-lactation, unmilking for 3 days, #5; and circulating leukocytes, #6]. Membranes were hybridized with [ $\alpha$ - $^{32}$ P]dCTP labeled cDNA inserts for bovine lactoferrin (pLf3a),  $\alpha_{s1}$ -casein (pC184), beta-casein (pC14), kappa-casein (pC371), beta-lactoglobulin (pB113), or alpha-lactalbumin (pALac), as indicated.

an mRNA for bLf of about 2500 bases, favorably agreeing with the 2351 base mRNA sequence which is devoid of a normal poly(A) tail (with 100 to 300 A's). Hybridization of liver RNA with bLf cDNA followed by low stringency washing gave a faint band at about 2500 bases. This band disappeared with high stringency washing and probably represents the homologous transferrin RNA.

The Northern blots (Figure 4) reflect relative changes in mRNA levels for each of the milk proteins at the represented mammary stages. Interestingly, little lactoferrin mRNA was noted in mammary tissue from either two days prepartum or lactation, even though the concentration of bLf in prelactating mammary secretion is generally 10-100 fold higher than in milk (8). This unexpected finding could be the result of translational regulation of bLf expression, or of protein storage and delayed release. The cessation of milking causes a dramatic increase in bLf mRNA which appears partly due to local non-hormonal effects, as evidenced by the increase in bLf mRNA in the



involuting unmilked half of the mammary gland of cow #3148 (Fig. 4, lane 4) while in the regularly milked contralateral half bLf mRNA remained low (Fig. 4, lane 3). The greater abundance of bLf mRNA in the fully unmilked half of the mammary gland of cow #3148 suggests increased induction of bLf in mammary tissue allowed to involute without continued milking stimulus. Although lactoferrin is found in secondary granules of polymorphonuclear leukocytes (3), no bLf mRNA was detected in RNA from circulating bovine peripheral leukocytes (Fig. 4, lane 6) either because the number of neutrophils in peripheral blood is low, or the abundance of bLf mRNA in circulating cells is low.

The half-life of bLf mRNA may be regulated in-part through the AUUUA and AUUUUA sequences found in the 3'UTR, as occurs with the destabilizing effect of similar sequences in the c-myc mRNA (32,33), by altering accessibility to endonuclease sensitive sites. Human lactoferrin does not have similar sequences which may explain the difference in milk secretion between human and bovine lactoferrins. Additional regulation of bLf expression may involve the formation of secondary structures in the mRNA as predicted for bases in the 5 to 305 region, and by the suboptimal sequence at the initiation codon.

Steady-state levels of mRNAs for  $\alpha_{s1}$ - and kappa- caseins did not decrease while those for beta-casein and beta-lactoglobulin decreased somewhat in the unmilked involuting half of the mammary gland of #3148 compared to the lactating contralateral half (Fig. 4, lane 4 and 3). Except for beta-lactoglobulin, mRNA levels for caseins and alpha-lactalbumin are significantly lower in early involuting mammary tissue (cow #2949) which was completely unmilked for three days prior to slaughter. Alpha-lactalbumin message was not detectable in the two day prepartum gland (Fig. 4 lane 1), and declined dramatically with early involution in the totally unmilked gland (Fig. 4 lane 5) compared to levels in lactating tissue (Fig. 4, lane 2), which were essentially the same as in the milked and unmilked halves of the same gland (Fig. 4, lanes 3 and 4), indicating a dramatic down regulation of bovine alpha-lactalbumin message soon after the complete cessation of milking.

These studies provide the first complete mRNA sequence for bovine lactoferrin with minor sequence differences from the two published partial sequences (18,19). The Northern blot data show for the first time that bovine mammary lactoferrin mRNA concentrations are regulated contrary to that for the other milk proteins. The exact mechanisms by which bLf mRNA is regulated must be experimentally determined.

**Acknowledgments:** We thank Dr. A.G. Mackinlay (U. New S. Wales, Kensington, Australia) for providing the casein and beta-lactoglobulin cDNAs and Dr. K. Brew for providing alpha-lactalbumin cDNA. We also thank Renee Fisk for performing the Northern blots and Ron Neiswander for assistance in preparation of the figures. Salaries and support provided by state and federal funds to the Ohio Agricultural Res. and Dev. Center, The Ohio State University; Manuscript No. 62-91.

#### REFERENCES

- (1) Rejman, J. J., Hegarty, H. M., & Hurley, W.L. (1989) *Comp. Biochem. Physiol.* 93B:929-934.
- (2) Crichton, R. R. (1990) *Advances in Protein Chem.* 40:284-263.
- (3) Masson, P. L. (1970) *La lactoferrine, proteine des secretions externes et des leucocytes neutrophiles* (Arsca ed.), Bruxelles.

- (4) Smith, K. L., & Schanbacher, F. L. (1977) *JAVMA* 170(10):1224-1469.
- (5) Weinberg, E. D. (1974). *Science* 184:952-956.
- (6) Nishiya, K. & Horwitz, D. A. (1982) *J. Immun.* 129(6):2519-2523.
- (7) Nichols, B. L., McKee, K. S., Henry, J. F., & Putman, M. (1987) *Pediat. Res.* 21(6):563-567.
- (8) Welty, F. K., Smith, K. L., & Schanbacher, F. L. (1976). *J. Dairy Sci.* 59:224-231.
- (9) Harmon, R. J., Schanbacher, F. L., Ferguson, L. C. and Smith, K. L. (1976) *Infect. Immun.* 13:533-542.
- (10) Lewis-Jones, D. I., Lewis-Jones, M. S., Connolly, R. C., Loyd, D. C., & West, C. R. (1985) *Pediatr. Res.* 19:561-565.
- (11) Yang, F., Lum, J. B., McGill, J. R. Moore, C. M., Naylor, S. L., Van Bragt, P., H., Baldwin, W. D., & Bowman, B. H.
- (12) Jeltsch, J.-M., & Chambon, P. (1982) *Eur. J. Biochem.* 122:291-295.
- (13) Powell, M. J., & Ogden, J. E. (1990) *Nuc. Acids Res.* 18:4013.
- (14) Rey, M. W., Woloshuk, S. L., de Boer, H. A. & Piper, F. R. (1990) *Nuc. Acids Res.* 18:5288.
- (15) Pentecost, B. T., & Teng, C. T. (1987) *J. Biol. Chem.* 262:10134-10139.
- (16) Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W. & Baker, E. N. (1989) *J. Mol. Biol.* 209:711-734.
- (17) Rochard, E., Legrand, D., Mazurier, J., Montreuil, J., and Spik, G. (1989) *FEBS Letters* 255:201-204.
- (18) Mead, P. E., & Tweedie, J. W., (1990) *Nuc. Acids Res.* 18:7167.
- (19) Pierce, A., Colavizza, D., Benaissa, M., Maes, P., Tartar, A., Montreuil, J., and Spik, G. (1991) *Eur. J. Biochem.* 196:177-184.
- (20) Goodman, R. E. (1990) Bovine mammary lactoferrin: cDNA cloning, northern blots, and analysis of the mRNA sequence and deduced protein structure. Thesis, Ohio State University.
- (21) Talhouk, R. S., Neiswander, R. L., and Schanbacher, F. L. (1990) *Tissue and Cell* 22(5):583-599.
- (22) Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular cloning* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- (23) Hahn, C. S., Strauss, E. G., and Strauss, J. H. (1989) *Methods Enzymol.* 180:121-130.
- (24) Puissant, C., and Houdebine, L.-M. (1990) *BioTechniques* 8(2):148-149.
- (25) Siegel, L. Il, and Bresnick, E. (1986) *Anal. Biochem.* 159:82-87.
- (26) Stewart, A. F., Willis, I. M., and Mackinlay, A. G. (1984) *Nucl. Acids Res.* 12(9):3895-3907.
- (27) Stewart, A. F., Bonsing, J., Beattie, C. W., Shah, F., Willis, I. M., and Mackinlay, A. G. (1987) *Mol. Biol. Evol.* 4(3):231-241.
- (28) Willis, I. M., Stewart, A. F., Caputo, A., Thompson, A. R., and Mackinlay, A. G. (1982) *DNA* 1(4):375-386.
- (29) Calzone, F. J., Britten, R. J., and Davidson, E. H. (1987) *Methods Enzymol.* 152:611-632.
- (30) Metz-Boutigue, M.-H., Jolles, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Montreuil, J., and Jolles, P. (1984) *Eur. J. Biochem.* 145:659-676.
- (31) Baldwin, G. S., and Weinstock, J. (1988) *Nuc. Acids Res.* 16:8720.
- (32) Malter, J. S. (1989) *Science* 246:664-666.
- (33) Wreschner, D. H., and Rechavi, G. (1988) *Eur. J. Biochem.* 172:333-340.
- (34) Kozak, M. (1987) *Nuc. Acids Res.* 15:8125-8149.
- (35) Rose, R. M., Plowman, G. D., Teplow, D. B., Dreyer, W. J., Hellstrom, K. E., and Brown, J. P. (1986) *Proc. Natl. Acad. Sci. USA* 83:1261-1265.
- (36) von Heijne, G. (1983) *Eur. J. Biochem.* 133:17-21.
- (37) Spik, G., Coddeville, B., and Montreuil, J. (1988) *Biochimie* 70:1459-1469.