Expression of cloned human lactoferrin in baby-hamster kidney cells

Kathryn M. STOWELL,* Thomas A. RADO,† Walter D. FUNK‡ and John W. TWEEDIE*§

*Department of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand, †Department of Medicine, Division of Hematology/Oncology, University of Alabama at Birmingham, AL, U.S.A., and ‡Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

Human lactoferrin was expressed from a cloned cDNA introduced into mammalian cells in tissue culture. Total RNA was extracted from human bone marrow, and lactoferrin cDNA was synthesized by primer-specific polymerase chain reaction after oligo(dT)-primed first-strand synthesis. The cDNA was sequenced to confirm its identity with previously published human lactoferrin sequences and cloned into the eukaryotic expression vector pNUT. Recombinant vector DNA containing the human lactoferrin sequence was introduced into baby-hamster kidney (BHK) cells in culture, and stable transfectants were produced by dominant marker selection. Human lactoferrin was expressed from the metallothionein promoter of pNUT by Zn²⁺ induction. The protein was secreted into the tissue-culture medium and was subsequently purified to homogeneity in a single step. Initial characterization suggests that the protein expressed by BHK cells is identical with native human lactoferrin.

INTRODUCTION

Lactoferrin is a member of the transferrin family of proteins which include serum transferrin, ovotransferrin and melano-transferrin (Aisen & Listowsky, 1980; Chasteen, 1983; Brock, 1985; Rose *et al.*, 1986). All of the transferrins are glycoproteins consisting of a single polypeptide chain of 650–700 amino acid residues. There is a 40–50% amino acid sequence identity between members of the family (Metz-Boutigue *et al.*, 1984). Transferrins also exhibit an amino acid sequence identity of ~40% between the *N*- and *C*-terminal halves of the protein, which is reflected in a characteristic bilobal tertiary structure and in the capacity to bind two Fe³⁺ ions (Gorinsky *et al.*, 1979; Metz-Boutigue *et al.*, 1984; Anderson *et al.*, 1987). Each lobe contains one iron-binding site, which binds a single Fe³⁺ ion concomitantly with one CO₃²⁻ anion (Anderson *et al.*, 1989). Sequence comparisons suggest that the iron ligands are the same for all transferrins, in both the *N*-terminal and *C*-terminal lobes, with the exception of melanotransferrin. This protein appears to have an intact iron-binding site in the *N*-terminal lobe, but changes which may prevent iron-binding in the *C*-terminal lobe (Baker *et al.*, 1987).

The functional role of the transferrins is undoubtedly related to their capacity for tight ($K_{app} \sim 10^{20}$), but reversible, binding of iron (Aisen & Listowsky, 1980). The serum transferrins have been shown to have a major function in the transport of iron (Crichton, 1985), and a receptor-mediated mechanism for the internalization of iron-loaded transferrin has been well characterized (Huebers & Finch, 1987). No clear role has been demonstrated for lactoferrin, but several functions have been suggested, including bacteriostasis (Reiter, 1983), iron nutrition (Brock, 1980), regulation of myelopoiesis (Broxmeyer *et al.*, 1978) and modulation of the inflammatory response (Oseas *et al.*, 1981).

The amino acid sequence of human lactoferrin has been known for some time (Metz-Boutigue *et al.*, 1984). The tertiary structure has been determined to 0.32 nm (3.2 Å) by X-ray-diffraction analysis of crystals of iron-loaded lactoferrin (Anderson *et al.*, 1987), and the structure has recently been refined to 0.28 nm (2.8 Å) resolution (Anderson *et al.*, 1989). Crystal-structure analysis of the iron-free form of lactoferrin has demonstrated a substantial conformational change that accompanies iron-binding and release (Anderson *et al.*, 1990). This information provides a detailed picture of the protein structure and introduces the possibility of altering the protein by site-directed mutagenesis of cloned copy DNA and examining the structural features which contribute to the ability of lactoferrin to bind iron. Site-directed mutagenesis can also be used to study possible protein–receptor interactions of lactoferrin. The ability to manipulate the protein to change these interactions will provide an experimental system to probe their role in the function of the protein.

As the first step in this investigation we synthesized a cDNA for human lactoferrin from bone-marrow RNA and cloned it into a eukaryotic expression vector. We have obtained expression of the protein in mammalian cells in tissue culture at levels sufficient to enable functional and crystallographic studies of the protein to be carried out. The recombinant protein appears to be identical with native milk lactoferrin by a number of criteria, suggesting that the system will be suitable for use in the expression of specific mutant lactoferrins.

EXPERIMENTAL

Materials

All restriction endonucleases, T₄ polynucleotide kinase, *Escherichia coli* DNA polymerase I (Klenow fragment), T₄ DNA

Abbreviations used: AMV, avian myeloblastosis virus; MT, metallothionein; BHK, baby-hamster kidney; DMEM, Dulbecco's modification of Eagle's medium; FCS, foetal-calf serum; PNGase, peptide:N-glycosidase.

§ To whom correspondence should be addressed.

ligase and oligo(dT) were obtained from Bethesda Research Laboratories (MD, U.S.A.). Avian-myeloblastosis-virus (AMV) reverse transcriptase and calf alkaline phosphatase were from Boehringer Mannheim. The strain of *E. coli* used for maintenance of all plasmids and of phage m13 was XL-1 (Stratagene, La Jolla, CA, U.S.A.). Polymerase-chain-reaction (PCR)-primer oligonucleotides were obtained from the Department of Immunobiology, University of Auckland Medical School, Auckland, New Zealand.

Isolation of neutrophil precursor cells

Neutrophil precursor cells were isolated by differential sedimentation through Ficoll-Paque (Pharmacia LKB Biotechnology, Uppsala, Sweden) from 30 ml of heparinized bone marrow obtained from an individual volunteer donor. The donor gave informed consent, and approval for the procedure was obtained from the Massey University Committee on Human Ethics. The cells were washed twice with balanced salt solution (0.126 M-NaCl/0.01 % glucose/5 μ M-CaCl₂/0.098 mM-MgCl₂/0.54 mM-KCl/0.0145 M-Tris, pH 7.6) and suspended in RNA extraction buffer. RNA was prepared immediately from these cells.

Preparation of RNA

The mononuclear cell fraction of bone marrow was homogenized in RNA extraction buffer (4.0 M-guanidine thiocyanate/0.025 M-citrate/0.7 % β -mercaptoethanol/0.5 % *N*-dodecanoylsarcosine, pH 7.0) with an Ultra-Turrax homogenizer (Chirgwin *et al.*, 1979). The homogenate was centrifuged through 5.7 M-CsCl/100 mM-EDTA, pH 7.0, at 33000 rev./min for 25 h in an SW41-Ti (Beckman) rotor. The RNA pellet was dissolved in extraction buffer and further purified by phenol/chloroform extraction, concentrated by precipitation with ethanol and resuspended in 10 mM-Hepes/1 mM-EDTA, pH 7.6.

Synthesis of cDNA

First-strand cDNA was prepared from 1 μ g of total cellular RNA by oligo(dT)-primed DNA synthesis with AMV reverse transcriptase as described by Sarkar & Sommer (1988). Double-stranded cDNA was prepared from the first-strand cDNA product by amplification using the PCR. Non-redundant oligonucleotide primers (25-mers) were designed using nucleotide sequences in the 5' and 3' untranslated regions of human lactoferrin cDNA (Rado *et al.*, 1987; T. A. Rado, K. M. Stowell & X.-P. Wei, unpublished work). The position of the 3' PCR primer was upstream of the polyadenylation signal for human lactoferrin cDNA. Transcription termination and polyadenylation signals were provided by vector sequences. The PCR primers were used without purification. PCRs were carried out by using 10 μ l of a 1:100 dilution of the products of first-strand cDNA synthesis, the specific primers at a final concentration of 0.1 μ M and other reagents as recommended by the supplier (Gene Amp; Cetus Corp., Norwalk, CT, U.S.A.) in a total volume of 100 μ l. In all, 30 cycles (94 °C, 1 min; 55 °C, 2 min; 72 °C, 2 min) were performed in a thermal cycler (Perkin-Elmer/Cetus).

Cloning the PCR product

The PCR reaction product was concentrated by ethanol precipitation and the 5'-OH groups phosphorylated using T₄ polynucleotide kinase. The human lactoferrin cDNA was then purified on a 1 % low-melting-point agarose gel and isolated using Gene-clean (Bio-101-Inc., La Jolla, CA, U.S.A.) after excision of the 2.3 kb band from the gel. The purified cDNA was

made blunt-ended using the Klenow fragment of *E. coli* DNA polymerase I and ligated into the restriction-endonuclease-*Hinc*II site of the vector pGEM-1 (Promega, Madison, WI, U.S.A.). After ligation, the recombinant vector was used to transform *E. coli*, and recombinant transformants were selected by colony hybridization using a partial human lactoferrin cDNA probe. The cDNA was excised from pGEM-1 using *Bam*HI and *Hind*III, made blunt-ended and subcloned into the *Sma*I site of m13mp18 for sequencing and into the *Sma*I site of pNUT (Palmiter *et al.*, 1987) to produce pNUT:hLF for transfection of BHK cells. The expression vector pNUT contains human growth-hormone genomic sequences which are under transcriptional control of the mouse metallothionein (MT-1) promoter. Cleavage of pNUT with *Sma*I releases most of the human-growth-hormone sequence, leaving only the 3' non-coding region, which contains transcription termination and polyadenylation signals (Funk *et al.*, 1990). All vectors were dephosphorylated using calf intestine alkaline phosphatase after cleavage by the appropriate restriction endonucleases.

DNA sequence analysis

Both strands of the cloned 2.3 kb PCR product were sequenced in m13mp18 using Sequenase version 2.0 (United States Biochemical Corp., Cleveland, OH, U.S.A.) after generating a series of nested deletions with Erase-a-base (Promega). The cloning site junction in pNUT:hLF was sequenced by double stranded sequencing with Sequenase version 2.0 using a primer (MT-30) specific for the metallothionein promoter of pNUT. The MT-30 primer and the vector pNUT were the gift of Dr. R. T. A. MacGillivray (Department of Biochemistry, University of British Columbia).

Maintenance of BHK cells

Baby hamster kidney (BHK) cells were obtained from Dr. R. T. A. MacGillivray. These were maintained on Dulbecco's modification of Eagle's medium (DMEM; Gibco, Grand Island, NY, U.S.A.), supplemented with 10 % (v/v) foetal-calf serum (FCS; Gibco) and the antibiotics streptomycin (100 μ g/ml) and penicillin (100 units/ml) (Sigma). The cells were cultured in a humidified atmosphere of air/CO₂ (19:1) at 37 °C. The medium was changed every 2 or 3 days and cells were passed at between 40 and 80 % confluence. Versene (0.5 mM-EDTA/0.14 M-NaCl/0.27 M-KCl/10 mM-Na₂HPO₄/1 mM-KH₂PO₄, pH 7.2) was used to release the adherent cells from the culture surface.

Transfection and expression

pNUT:hLF plasmid DNA which had been purified by CsCl gradient ultracentrifugation (Sambrook *et al.*, 1989) was introduced into BHK cells as a calcium phosphate co-precipitate using a Cell Pfect kit (Pharmacia LKB Biotechnology, Uppsala, Sweden). Transfectants were selected and maintained on DMEM/FCS containing methotrexate (0.5 mM).

Expression of lactoferrin

Stable transfectants were grown to near-confluence in 75 cm² tissue-culture flasks (Nunc, Roskilde, Denmark). Expression of lactoferrin from the MT-1 promoter of pNUT was induced by the addition of 80 μ M-ZnSO₄. Medium was collected every day for 4 weeks and pooled. Lactoferrin was identified in the medium by immunoprecipitation with antibodies against human milk lactoferrin.

Preparation of antibodies

Polyclonal antibodies against human lactoferrin were raised in New Zealand White rabbits by standard protocols (Crowle,

1973) and purified by affinity chromatography using human lactoferrin-Sepharose prepared as described by Bethel *et al.* (1979).

Purification of lactoferrin

Recombinant lactoferrin was purified from culture medium by ion-exchange chromatography on CM-Sephadex C-50 (Pharmacia LKB Biotechnology, Uppsala, Sweden) equilibrated with 0.025 M-Tris(pH 7.5)/0.1 M-NaCl. The culture medium was centrifuged at 10000 *g* for 10 min to remove any cell debris, and the supernatant was loaded directly on to the equilibrated CM-Sephadex. After washing with equilibration buffer the lactoferrin was eluted with a salt gradient from 0.1 to 1.1 M-NaCl. The fractions containing lactoferrin were identified by immunoprecipitation. The purity of the protein in these fractions was determined by SDS/PAGE. Fractions containing lactoferrin were pooled, concentrated, then dialysed against 0.025 M-Tris (pH 7.8)/0.2 M-NaCl/0.01 M-NaHCO₃.

SDS/PAGE

Purified lactoferrin and lactoferrin immunoprecipitates were analysed by SDS/PAGE as described by Laemmli (1970) using the Ornstein (1964) and Davis (1964) buffer system with the addition of 0.1 % SDS in gels and electrophoresis buffer. Gels were stained with Coomassie Brilliant Blue R.

RESULTS

Analysis of cDNA

The 2.3 kb cDNA produced by PCR was analysed by digestion with restriction endonucleases diagnostic for human lactoferrin cDNA. The restriction fragments produced were the sizes predicted from the nucleic acid sequence of human neutrophil lactoferrin (Rado *et al.*, 1987; T. A. Rado, K. M. Stowell & X.-P. Wei, unpublished work). The nucleic acid sequence of the 2.3 kb PCR product included the entire coding sequence of human lactoferrin, including the 19-residue leader sequence plus 17 nucleotides of the 5' untranslated region and 108 nucleotides of the 3' untranslated region. Comparison of this sequence with that determined by T. A. Rado, K. M. Stowell & X.-P. Wei (unpublished work) demonstrated two changes which would be reflected in the amino acid sequence of the recombinant protein and two changes in codon usage which would not alter the amino acid sequence. There is an additional arginine residue at position 4 of the mature protein and an arginine substitution for a lysine residue at position 28. The substitution at position 28 is the result of a single nucleotide change. In terms of protein structure these are both conservative changes, with the arginine-4 being in a region with four contiguous arginine residues. It is noteworthy that the original amino acid sequence of human milk lactoferrin (Metz-Boutigue *et al.*, 1984) demonstrated four arginine residues in this part of the protein sequence. The cDNA sequence of human mammary lactoferrin has recently been published (Powell & Ogden, 1990). This sequence indicates the presence of four arginines near the *N*-terminus of the mature protein and a lysine at position 28. Arginine-4 and arginine-28 in our sequence are in regions of the protein which lie on the outside of the molecule (Anderson *et al.*, 1989) and therefore would not be expected to influence protein folding or iron binding. The 'silent' changes in our cDNA occurred at nucleotide positions 1089 (T substituted instead of C) and 1245 (G substituted instead of A) of the sequence reported by Powell & Ogden (1990). Apart from the difference noted (lysine-28 to arginine-28), these results suggest

Fig. 1. SDS/PAGE of culture medium and immunoprecipitates

Recombinant lactoferrin was immunoprecipitated from culture medium by the addition of affinity-purified antibodies specific for human lactoferrin. Immunoprecipitates were formed over a sucrose step gradient (0.5 ml of 0.5 M- and 0.5 ml of 1.0 M-sucrose) in the presence of 0.1 % Triton X-100 and 0.1 % sodium deoxycholate in phosphate-buffered saline (0.01 M-phosphate/0.15 M-NaCl, pH 7.2) and were collected by centrifugation for 5 min at 12000 rev./min (r_{av} 4.5 cm) in a Microfuge. Immunoprecipitates were washed three times with 0.1 % Triton X-100/0.1 % sodium deoxycholate in phosphate-buffered saline. SDS/polyacrylamide gels (8% acrylamide) were prepared, run and stained as described in the Experimental section. 1, A 5 μ l portion of culture medium from pNUT:hLF-transfected BHK cells; 2, immunoprecipitate from 200 μ l of culture medium from pNUT:hLF-transfected BHK cells; 3, 5 μ l of culture medium from untransfected BHK cells; 4, immunoprecipitate from 200 μ l of culture medium from untransfected BHK cells; 5, affinity-purified anti-(human lactoferrin) immunoglobulins; 6, lactoferrin purified from human milk.

that the amino acid sequence of mammary lactoferrin is identical with that of neutrophil lactoferrin.

Sequence analysis of the cloning site junctions in pNUT:hLF demonstrated that no nucleotides had been deleted or added to these regions during cloning.

Identification of human lactoferrin expressed in BHK cells

Lactoferrin immunoprecipitated from the culture medium had a mobility on SDS/polyacrylamide gels which was identical with that of lactoferrin isolated from human milk (Fig. 1). Untransfected BHK cells did not secrete any protein which was immunoprecipitable by antibodies to human lactoferrin.

Purification of recombinant human lactoferrin

Recombinant human lactoferrin was eluted from CM-Sephadex (Fig. 2) as a single symmetrical peak, as shown by A_{280} . Fractions from this peak showed a single Coomassie Blue-staining band on SDS/PAGE. The concentration of lactoferrin in the culture medium was estimated to be 20 mg/litre by determination of the A_{280} of the pooled column fractions.

Characterization of recombinant human lactoferrin

Absorption spectrum. The absorption spectrum of recombinant human lactoferrin is shown in Fig. 3 and is identical with that of

Fig. 2. Purification of recombinant human lactoferrin

Elution profile of recombinant human lactoferrin from CM-Sephadex C-50. ●, A_{280} ; ○, [NaCl] in elution buffer.

Fig. 3. Absorption spectrum of recombinant human lactoferrin

The absorption spectrum of a solution [1.34 mg/ml in 0.025 M-Tris/HCl (pH 7.5)/0.2 M-NaCl/0.01 M-NaHCO₃] of the recombinant protein was determined using a Hewlett-Packard diode-array spectrophotometer (HP 8452A). (a) Absorbance scale 0–1.5; (b) Absorbance scale 0–0.15.

the native milk protein between 260 nm and 700 nm. The A_{280}/A_{466} ratio of the recombinant proteins was 22 and the A_{410}/A_{466} ratio was 0.85. This indicates that the protein was fully saturated with iron and was confirmation of its purity (Parry & Brown, 1974).

N-Terminal amino acid sequence. The recombinant lactoferrin was sequenced from the *N*-terminus through 21 residues and the sequence was identical with that predicted from the cDNA sequence. The *N*-terminus of the protein sequence corresponds to the *N*-terminus of secreted milk lactoferrin (Metz-Boutigue *et al.*, 1984). This result shows that the pre-protein was processed correctly during synthesis and secretion from BHK cells. The presence of four arginine residues at the *N*-terminus of the recombinant protein confirms the prediction from the sequence of the cDNA derived from bone-marrow cells. This result is at variance with the sequence of a cDNA clone derived from white blood cells obtained from a patient with chronic myeloid leukaemia (Rado *et al.*, 1987; T. A. Rado, K. M. Stowell & X.-P. Wei, unpublished work). The protein sequence of human milk lactoferrin (Metz-Boutigue *et al.*, 1984) and the cDNA sequence of a clone isolated from a human mammary-gland library (Powell & Ogden, 1990) have four arginine residues at this position. These variations in sequence prompted us to determine the *N*-terminal sequence of a sample of human milk lactoferrin. This sample, from an individual donor, showed a single arginine in this region of the protein, with an otherwise identical *N*-terminal amino acid sequence (Table 1). This result suggests the possibility of some variation between different individuals in this region of the protein in the amino acid sequence of lactoferrin.

Deglycosylation of lactoferrin. Human milk lactoferrin and recombinant human lactoferrin produced by BHK cells were deglycosylated under identical conditions using the endoglycosidase peptide:*N*-glycosidase (PNGase). After a 24 h incubation with PNGase, the human milk lactoferrin appeared to be completely deglycosylated, as judged by a decrease in apparent molecular mass on SDS/polyacrylamide gels (Fig. 4). Most of the recombinant human lactoferrin demonstrated an identical shift in apparent molecular mass, but a small fraction of the protein was resistant to deglycosylation by PNGase. This observation, plus the observation that the recombinant protein shows a more indistinct band on SDS/polyacrylamide gels, suggests a possible heterogeneity in glycosylation of the protein produced in BHK cells.

pH-dependent release of iron from lactoferrin. The pH dependence of iron release from recombinant human lactoferrin was identical to that of human milk lactoferrin (Fig. 5). Both

Table 1. Sequence diversity of the *N*-terminus of human lactoferrin

The *N*-terminal amino acid sequence of recombinant human lactoferrin and human milk lactoferrin was determined by the automated Edman degradation procedure using an Applied Biosystems gas-phase sequencer (model 470A). The cDNA sequence of recombinant human lactoferrin was determined as described in the Experimental section and 'translated' into the amino acid sequence shown.

Protein	Sequence	Reference
Recombinant protein	Gly-Arg-Arg-Arg-Arg-Ser-Val-Gln-Trp-Xaa-Ala-Val-	The present study
cDNA (neutrophil)	Gly-Arg-Arg-Arg-Arg-Ser-Val-Gln-Trp-Cys-Ala-Val-	The present study
Milk protein	Gly-Arg- ... - ... -Ser-Val-Gln-Trp-Xaa-Ala-Val-	The present study
cDNA (neutrophil)	Gly-Arg-Arg-Arg-Arg-Ser-Val-Gln-Trp-Cys-Ala-Val-	Metz Boutigue <i>et al.</i> (1984)
	Gly-Arg-Arg- ... -Arg-Ser-Val-Gln-Trp-Cys-Ala-Val-	T. A. Rado, K. M. Stowell & X.-P. Wei, unpublished work
cDNA (mammary gland)	Gly-Arg-Arg-Arg-Arg-Ser-Val-Gln-Trp-Cys-Ala-Val-	Powell & Ogden (1990)

Fig. 4. SDS/PAGE of glycosylated and deglycosylated lactoferrin

Lactoferrin ($2 \mu\text{g}/\mu\text{l}$) was deglycosylated using partially purified *Flavobacterium meningosepticum* PNGase (peptide: *N*-glycosidase) ($\sim 1 \mu\text{g}/\mu\text{l}$) by incubation overnight at room temperature in 0.1 M-phosphate (pH 6.5)/0.05 M-EDTA and 0.5% Nonidet P40. SDS/polyacrylamide gels (8%) were prepared, run and stained as described in the Experimental section. 1, PNGase; 2, recombinant human lactoferrin ($\sim 5 \mu\text{g}$) treated with PNGase; 3, recombinant human lactoferrin ($\sim 5 \mu\text{g}$); 4, human milk lactoferrin ($\sim 5 \mu\text{g}$) treated with PNGase; 5, human milk lactoferrin ($\sim 5 \mu\text{g}$); 6, M_r markers (phosphorylase *b*, 97.4 kDa; BSA, 66 kDa; ovalbumin, 45 kDa; carbonic anhydrase, 29 kDa). The protein band at M_r 36000 in lanes 1, 2, and 4 corresponds to PNGase. The band at M_r 45000 in these lanes is an unknown contaminant of the PNGase preparation.

Fig. 5. pH-dependent release of iron from lactoferrin

The absorbance of solutions of iron-saturated human recombinant and milk lactoferrin was measured at 466 nm after dialysis against buffers of the pH values indicated. ●, Recombinant lactoferrin; ○, milk lactoferrin. The buffers used for the various pH ranges were: pH 8–7, 0.05 M-Tris/HCl; pH 6.5–6, 0.05 M-Mes; pH 5.5–3.5, sodium acetate (I 0.05); pH 3–2, 0.1 M-glycine/HCl. All buffers contained 0.2 M-NaCl.

proteins began to lose the iron at around pH 4 and appeared to be almost totally iron-free at pH 2.0. This result suggests that the iron environment of the recombinant protein is the same as that of milk lactoferrin and that the factors involved in the release of iron with decreasing pH are the same for the two proteins.

DISCUSSION

Human lactoferrin cDNA has been produced from RNA using PCR, cloned and expressed in a eukaryotic expression system. The protein has been purified to homogeneity from the culture medium by a single ion-exchange-chromatographic step. Initial characterization of the protein suggests that it is identical with lactoferrin from human milk.

Sequence analysis has revealed a number of differences between the nucleotide sequence of this cDNA and that previously reported for human neutrophil (Rado *et al.*, 1987; T. A. Rado, K. M. Stowell & X.-P. Wei, unpublished work) and human mammary gland (Powell & Ogden, 1990) lactoferrin cDNA. The previously reported sequences were determined from analysis of cDNA clones isolated without PCR amplification from conventional cDNA libraries. The error rate of *Taq* polymerase during DNA replication under PCR conditions is reported to be ~ 1 in 9000 bases for single base substitutions and to be 1 in 41000 for frameshift errors (Tindall & Kunkel, 1988). We feel that it is unlikely that the PCR would introduce a complete codon for arginine. It is worth noting that the additional arginine codon in our cDNA is different from the two flanking arginine codons and that the sequence in this region is identical with that recently reported by Powell & Ogden (1990). The PCR is also unlikely to produce such a conservative change as lysine to arginine at position 28, although this change has resulted from a single nucleotide substitution. The residue at position 28 of both previously reported human lactoferrin cDNAs is lysine. PCR has been reported to add non-template directed nucleotides at 3' ends (Denney & Weissmann, 1990). However, DNA sequence analysis of the PCR generated cDNA confirmed that the ends were identical with the sequence predicted by the oligonucleotide primers. We suggest that the conservative nature of the observed changes and the alterations in codon usage shown by our cDNA are most likely to be due to allelic differences frequently observed between individuals. However, the possibility that this sequence diversity is caused by *Taq* polymerase cannot be ruled out, particularly in the case of the lysine-to-arginine substitution at position 28. Analysis of selectively amplified regions of genomic DNA from a number of individuals should demonstrate conclusively whether these differences are allelic or PCR-induced. The most important point is that the amino acid sequence predicted by our cDNA sequence should give rise to a recombinant protein that is identical in structure and function with native lactoferrin.

Recombinant lactoferrin, either purified or in immunoprecipitates from culture medium, co-migrates with milk lactoferrin on SDS/polyacrylamide gels. This suggests that the two proteins have identical molecular masses. However, the recombinant protein band is always slightly less sharp than the corresponding milk lactoferrin band, suggesting some type of heterogeneity in the recombinant protein. Analysis of deglycosylated proteins showed that, whereas there was complete deglycosylation of the milk protein, a fraction of the recombinant protein remained resistant to deglycosylation after an extended incubation with PNGase. The completely deglycosylated proteins co-migrated on SDS/polyacrylamide gels, and both deglycosylated milk and recombinant lactoferrins were present as sharp bands. This result suggests that the proteins have the same molecular mass, but that there may be some glycosylation differences. BHK cells have been reported to glycosylate recombinant proteins in a manner different from that occurring in the tissue in which synthesis normally occurs (Tsuda *et al.*, 1988). BHK cells are reported to form tri- and tetra-antennary structures which may not be recognized by PNGase (Plummer *et al.*, 1984; Steube *et al.*, 1986). Comparison of the glycan moieties of human

milk and neutrophil lactoferrin suggests that there are tissue-specific differences in the glycosylation of lactoferrin (Spik *et al.*, 1988). Differences in glycosylation should not cause problems in X-ray-crystallographic analysis, as the carbohydrate moieties have a very indistinct electron density, suggesting a high degree of mobility in the protein crystal (Anderson *et al.*, 1989). Should the carbohydrate interfere with crystallization it will be possible to prepare the deglycosylated protein by the action of endoglycosidases, as shown here, or by growing tissue-culture cells expressing lactoferrin in the presence of tunicamycin (Heifetz *et al.*, 1979). The nature of the carbohydrate side chains of transferrins has been implicated in receptor recognition (Prieels *et al.*, 1978; Davidson & Lonnerdal, 1988). However, other reports suggest that the carbohydrate may not be involved (Hu *et al.*, 1988; Mazurier *et al.*, 1989). If in fact the carbohydrate structure is critical for receptor recognition and binding, then it may be necessary to explore other host cells for expression of recombinant lactoferrin before recombinant protein can be used for studying receptor-protein interactions.

The visible absorption spectrum of recombinant human lactoferrin is identical with that of the milk protein. This suggests that the iron-binding environments of the two proteins are identical. The recombinant protein purified from the culture medium is completely iron-saturated. The fact that the recombinant protein is capable of this degree of iron saturation indicates that the iron environment, and hence protein folding, are probably identical with that of the milk protein. This conclusion is also supported by the observation that the pH-dependence of the release of iron from recombinant protein is identical with that of iron release from milk protein.

The result of *N*-terminal sequence analysis of the recombinant protein indicates that correct processing of the signal peptide (present in the cDNA) has occurred during transfer of recombinant lactoferrin through the endoplasmic reticulum of BHK cells. The significance of the observed *N*-terminal sequence diversity between both milk lactoferrins and neutrophil lactoferrin remains to be clarified. This part of the protein is on the outside of the tertiary structure, and the electron density attributable to the *N*-terminal region of the protein is very indistinct (Anderson *et al.*, 1989). This suggests that the extreme *N*-terminal region is not involved in protein folding, and any diversity in this region is unlikely to lead to any functional differences. A more extensive investigation of the *N*-terminal sequence of milk lactoferrin from a number of individuals should demonstrate the extent of this diversity.

Crystallization and X-ray-diffraction analysis of the recombinant protein lactoferrin will ultimately demonstrate the identity of the recombinant protein with milk lactoferrin. Our preliminary characterization suggests that the proteins are the same.

The successful cloning and expression of human lactoferrin cDNA now provides the appropriate tools for a detailed study of the protein. Site-directed mutagenesis can be used to study specific regions of the protein implicated in iron-binding and release, and in particular in the large conformational change which occurs in the *N*-terminal half of the molecule during iron uptake and release (Anderson *et al.*, 1990). The system used here for culture of BHK cells containing the pNUT:hLf construct has not been optimized for lactoferrin synthesis. However, lactoferrin levels of 20 mg of purified protein/litre of culture medium have been obtained. This level of expression will provide sufficient quantities of recombinant protein to undertake structural and functional studies, including X-ray crystallography. The optimization of lactoferrin synthesis by clonal selection and the use of more sophisticated culture systems with a high surface-area/culture-medium ratio should enable the production of even

greater quantities of protein. It may be necessary to explore these systems for the production of mutant proteins which may not be expressed as well as the naturally occurring protein.

This work was supported in part by grants from the Massey University Research Fund, New Zealand University Grants Committee, Palmerston North Medical Research Foundation and the Medical Research Council of New Zealand. We thank Dr. S. Gibbons (Palmerston North Public Hospital) for performing the bone-marrow biopsy. We also thank Mrs. Carole E. Flyger and Miss Carolyn Young for expert technical assistance, Mrs. Heather Baker and Dr. Gill E. Norris for helpful advice concerning the purification and characterization of recombinant lactoferrin and for providing PNGase, Dr. Graeme G. Midwinter for performing the *N*-terminal amino acid sequence analysis and Mr. Paul E. Mead, Miss Catherine L. Day, Professor E. N. Baker and Dr. Graham G. Pritchard for their support and helpful criticism.

REFERENCES

- Aisen, P. & Listowsky, I. (1980) *Annu. Rev. Biochem.* **49**, 357-393
- Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M. & Baker, E. N. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1767-1773
- Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W. & Baker, E. N. (1989) *J. Mol. Biol.* **209**, 711-734
- Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V. & Baker, E. N. (1990) *Nature (London)* **344**, 784-787
- Baker, E. N., Rumball, S. V. & Anderson, B. F. (1987) *Trends Biochem. Sci.* **12**, 350-353
- Bethel, G. S., Ayers, J. S., Hancock, W. S. & Hearn, M. T. W. (1979) *J. Biol. Chem.* **254**, 2572-2574
- Brock, J. H. (1980) *Arch. Dis. Childhood* **55**, 417-421
- Brock, J. H. (1985) *Metalloproteins: Part II* (Harrison, P. M., ed.), pp. 183-262, Macmillan, London
- Broxmeyer, H. E., Smithyman, A., Eger, R. R., Meyers, P. A. & de Sousa, M. (1978) *J. Exp. Med.* **148**, 1052-1067
- Chasteen, N. D. (1983) *Adv. Inorg. Biochem.* **5**, 201-233
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294-5299
- Crichton, R. R. (1985) in *Proteins of Iron Storage and Transport* (Spik, G., Montreuil, J., Crichton, R. R. & Mazurier, J., eds.), pp. 99-110, Elsevier, Amsterdam
- Crowle, A. J. (1973) *Immunodiffusion*, 2nd edn., pp. 78-84, Academic Press, New York
- Davidson, L. A. & Lonnerdal, B. (1988) *Am. J. Physiol.* **254**, G580-G585
- Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404-427
- Denney, D., Jr. & Weissmann, I. (1990) *Amplifications* **4**, 25-26
- Funk, W. D., MacGillivray, R. T. A., Mason, A. B., Brown, S. A. & Woodworth, R. C. (1990) *Biochemistry* **29**, 1654-1660
- Gorinsky, B., Horsburgh, C., Lindley, P. F., Moss, D. S., Parkar, M. & Watson, J. L. (1979) *Nature (London)* **281**, 157-158
- Heifetz, A., Keenan, R. W. & Elbein, A. D. (1979) *Biochemistry* **18**, 2186-2192
- Huebers, H. A. & Finch, C. A. (1987) *Physiol. Rev.* **67**, 520-582
- Hu, W.-L., Mazurier, J., Sawatzki, G., Montreuil, J. & Spik, G. (1988) *Biochem. J.* **249**, 435-441
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680-685
- Mazurier, J., Legrand, D., Hu, W.-L., Montreuil, J. & Spik, G. (1989) *Eur. J. Biochem.* **179**, 481-487
- Metz-Boutigue, M.-H., Jolles, J., Mazurier, J., Schoentgen, F., LeGrand, D., Spik, G., Montreuil, J. & Jolles, P. (1984) *Eur. J. Biochem.* **145**, 659-676
- Ornstein, L. (1964) *Ann. N.Y. Acad. Sci.* **121**, 321-349
- Oseas, R., Yang, H.-H., Baehner, R. L. & Boxer, L. A. (1981) *Blood* **57**, 939-945
- Palmiter, R. D., Behringer, R. R., Quaife, C. J., Maxwell, F., Maxwell, I. H. & Brinster, R. L. (1987) *Cell (Cambridge, Mass.)* **50**, 435-443
- Parry, R. M. Jr. & Brown, E. M. (1974) *Adv. Exp. Med. Biol.* **48**, 141-160
- Plummer, T. H., Jr., Elder, J. H., Alexander, S., Phelan, A. W. & Tarentino, A. L. (1984) *J. Biol. Chem.* **259**, 10700-10704
- Powell, M. J. & Ogden, J. E. (1990) *Nucleic Acids Res.* **18**, 4013
- Prieels, J.-P., Pizzo, S. V., Glasgow, L. R., Paulson, J. C. & Hill, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2215-2219
- Rado, T. A., Wei, X. & Benz, E. J., Jr. (1987) *Blood* **70**, 989-993

- Reiter, B. (1983) *Int. J. Tissue Reactions* **5**, 87–96
- Rose, T. M., Plowman, G. D., Teplow, D. B., Dreyer, W. J., Hellstrom, K. E. & Brown, J. P. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1261–1265
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd edn., pp. 1.33–1.46, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sarkar, G. & Sommer, S. S. (1988) *Nucleic Acids Res.* **16**, 5197
- Spik, G., Coddeville, B. & Montreuil, J. (1988) *Biochimie* **70**, 1459–1469
- Steube, K., Gross, V., Hosel, W., Tran-Thi, T.-A., Decker, K. & Heinrich, P. C. (1986) *Glycoconjugate J.* **3**, 247–254
- Tindall, K. R. & Kunkel, T. A. (1988) *Biochemistry* **27**, 6008–6013
- Tsuda, E., Goto, M., Murakami, A., Akai, K., Ueda, M., Kawanishi, G., Takahashi, N., Sasaki, R., Chiba, H., Ishihara, H., Mori, M., Tejima, S., Endo, S. & Arata, Y. (1988) *Biochemistry* **27**, 5646–5654

Received 19 September 1990/26 November 1990; accepted 29 November 1990