J Clin Pathol: Mol Pathol 1996;49:M74–M79

Nomenclature for proteins: is calprotectin a proper name for the elusive myelomonocytic protein?

M K Fagerhol

Introduction

"New" proteins will continue to be discovered, and suitable names for them should be sought. In many instances preliminary symbols are introduced which serve a practical purpose until the proteins have been characterised and appropriate descriptive names can be chosen.

Using modern molecular biological methods, the existence of a protein may be suspected before it has ever been seen on an analytical gel. By cloning and gene manipulaton it may also be possible to produce proteins that may never appear in the human body. This raises a fundamental question: when should the specific protein be regarded as such and be given a proper name? When the DNA sequence is determined? When the specific mRNA is formed? When single polypeptides emerge from ribosomes? When glycosylated or otherwise modified polypeptides are detected within the cell? When complexes of more than one polypeptide chain are formed? Or when the functionally active substance is released?

The very fact that such questions have not yet been discussed in depth in the scientific literature, let alone been covered by international conventions, suggests that the problems involved are difficult to address.

In a time when biological sciences are becoming increasingly dominated by patents and their use for commercial purposes, it is not surprising that different scientific groups may be reluctant to agree that they are, in fact, working with the same protein: it might spoil the chance of obtaining a separate patent. In the case of calprotectin, which will be used as an example below, scientists have been unwilling to exchange reagents or use proper analytical methods to clarify whether various groups were working with the same protein.

There is a good tradition and sound basis for giving descriptive names for well characterised, functional proteins. Their genes should also be named accordingly. As characterisation and delineation of biological activities may take considerable time, the use of preliminary symbols or short notations for proteins will still be necessary.

Blood Bank and Department of Immunology, Ullevaal University Hospital, 0407 Oslo, Norway

Accepted for publication 8 August 1995

WHAT IS A PROTEIN?

Polypeptides can be subdivided into: oligopeptides (five amino acids); polypeptides (six to 40 amino acids); and proteins (>40 amino acids, corresponding to a molecular mass of about 5 kDa). Some proteins are more complex and may have carbohydrate side chains or consist of complexes of two or more polypeptide chains (oligomeric proteins). The term quaternary structure has been introduced for specific, complex oligomeric proteins. Individual polypeptide chains are called protomers, monomers or subunits, which combine to form homo- (one chain type only) and hetero-oligomers (different chain types encoded by different genes)—for example, heterodimers.

Oligomeric protein subunits are mainly joined by non-covalent bonds, although covalent, disulphide bonds occur between cysteine residues and may confer additional stability on the quaternary structure of the protein.

MATERIALS AND METHODS FOR STUDYING PROTEIN STRUCTURE

Degradation during purification of protein, for instance by lysosomal enzymes or enzymes involved in specific activation or inactivation, must be avoided.

Conventional sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE) and two dimensional electrophoresis using isoelectric focusing in the first dimension and SDS electrophoresis in the second, are standard procedures for determining the molecular mass and composition of oligomeric proteins. It must be remembered, however, that these procedures involve intentional splitting of oligomeric proteins. Often, a combination of heating in a boiling waterbath in the presence of SDS alone or together with a potent detergent, 9 M urea, a reducing agent, such as dithiothreitol, followed by alkylation to prevent formation of new covalent, disulphide bonds, is used. After such treatment, even the familiar heterotetramer IgG protein may have only two bands on SDS-PAGE. This method is clearly unsuitable for determining the quaternary structure of IgG. The information garnered from SDS-PAGE should be combined with data from gel filtration (gel permeation chromatography) or analytical ultracentrifugation of the native protein. Details of the configuration of proteins may be revealed by x ray crystallography or nuclear magnetic resonance imaging, but such methods are available only in a limited number of laboratories.

Using antibodies raised against the subunits, recombinant proteins or synthetic polypeptide analogs, it may be possible to detect subunit epitopes in biological fluids, cell smears or

tissue sections. However, monoclonal antibodies are not suitable for studying the molecular configuration of a protein. Staining with antibodies can show the presence of one or more epitopes in cells, but this does not provide any information about the size of the protein carrying the epitope or whether two or more epitopes are situated on the same macromolecule. A negative reaction with certain monoclonal antibodies may not necessarily mean that the relevant polypeptide is absent: epitopes may be lost due to partial proteolysis; certain epitopes may be hidden in the intact oligomeric protein, by complex formation with other proteins or macromolecules in the tissue, or by the fixation procedures. Fixation procedures, as well as details of immunohistochemical protocols, are often optimised so that it is possible to see what one wants to see, which consequently means that the whole truth is not revealed.

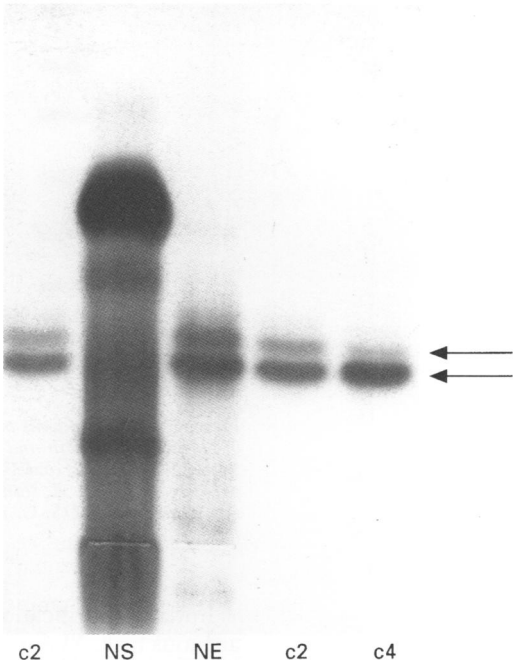

Figure 1 Protein band patterns of normal human serum (NS), crude neutrophil extracts (NE) and purified calprotectin, 2 or 4 g/l (c2, c4), after agarose gel electrophoresis with 75 mM sodium barbital buffer (pH 8·6) containing 2·5 mM EDTA-K_2. The calprotectin bands are indicated by the arrows.

Isolation and purification of calprotectin

Neutrophil granulocytes were isolated from standard units of citrated (CPD) blood from the blood bank as described previously,[1] except that EDTA-K_2 was added to a final concentration of 2·5 mM within one hour of venepuncture. The cells were resuspended in four volumes of 20 mM sodium barbital, 0·75 mM EDTA-K_2, 10 mM diisopropylfluorophosphate (DFP) (pH 7·5). The cytosol fraction was isolated as described by Dale *et al.*[2] Calprotectin was purified by anion exchange chromatography on a 2·5 × 10 cm column packed with DEAE–Sepharose FastFlow (Pharmacia, Uppsala, Sweden), equilibrated with 20 mM sodium barbital and 0·75 mM EDTA-K_2 (pH 7·5). After washing out non-bound proteins, calprotectin was eluted with 20 mM sodium barbital and 10 mM calcium chloride (pH 7·5). The eluate was concentrated by ultrafiltration using an immersible-CX filter (Millipore, Bedford, MA, USA). The concentrate was run on a 2·5 × 30 cm Sephadex G-25 column (Pharmacia) equilibrated with 50 mM ammonium bicarbonate, and then lyophilised. The dry protein dissolves rapidly in water or buffer.

Enzymatic degradation of calprotectin was then carried out by granulocyte derived lysosmal proteins. The granulocytes were isolated as described earlier, except that DFP was omitted.

The cells were disrupted[2] and centrifuged at 100 000 × *g* for 60 minutes. The lysosome rich sediment was harvested and resuspended in four volumes of 0·34 M sucrose, 20 mM imidazole, 2 mM ATP, 5 mM EGTA, 5 mM DTT, and 0·1% Triton X-100. The suspension was frozen and thawed three times, centrifuged at 100 000 × *g* for 60 minutes and the supernatant fluid harvested and used as the enzyme solution (E*). Mixtures of calprotectin, E* with or without calcium chloride or enzyme inhibitors (phenylmethylsulphonyl fluoride (PMSF) or DFP) or the antimicrobials sodium azide (final concentration 15 mM) or thimerosal (final concentration 0·25 mM), were incubated in a water bath at 37°C for 10 to 60 minutes. Enzymatic activity was subsequently inhibited by addition of DFP or PMSF (final concentration 5 to 10 mM) to each sample before quantitation of calprotectin by single radial immunodiffusion or analytical agarose gel elctrophoresis.[2] Concentrations are given as the per cent value at time t0.

Protein isolates were subjected to SDS electrophoresis on 8–18% polyacrylamide gels (ExcelGel, Pharmacia) using the LMW calibration kit (Pharmacia) which contains the following standards: phosphorylase b (94 kDa); bovine serum albumin (67 kDa); ovalbumin (43 kDa); carbonic anhydrase (30 kDa); soybean trypsin inhibitor (20·1 kDa); and α-lactalbumin (14·4 kDa). Before each run, 10 volumes of 1% SDS in 75 mM Tris-acetate buffer with 75 mM EDTA-K_2 and 0·01% bromphenol blue (pH 8·0) were added to each sample. The standards and samples, dissolved in the same buffer, were heated on a boiling water bath for five minutes, cooled on ice and 1/20 volumes of iodacetamide were added (40 mg dissolved in 1 ml of distilled water).

What is the relation between calprotectin, L1 protein, p8,14, MRP-8, MRP-14, calgranulins A and B, and cystic fibrosis antigen?

The existence of the protein now known as calprotectin was suspected in the late 1970s. The cytosol fraction of human neutrophil granulocytes gave a very prominent band on agarose gel electrophoresis (fig 1), and the protein was provisionally named L1 protein.[3] The name calprotectin was suggested when the protein was found to have antimicrobial properties.[4]

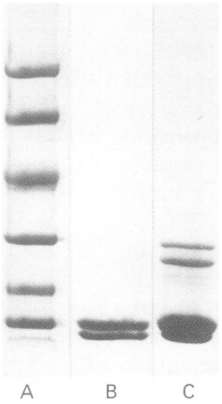

Figure 2 Band patterns produced on SDS electrophoresis of purified calprotectin and molecular mass standards of 94, 67, 43, 30, 20, and 14·4 kDa run on an 8–18% polyacrylamide gel. A, standards; B, calprotectin (0·5 mg/ml) treated with SDS; C, untreated calprotectin (3·5 mg/ml).

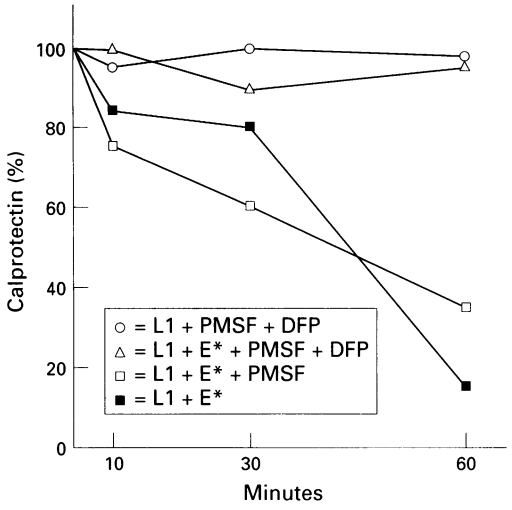

Figure 3 Degradation of calprotectin (L1) by lysosomal enzymes (E) from human granulocytes and the inhibitory effect of 10 mM PMSF and 5 mM DFP. Calprotectin concentrations are expressed as the per cent value at time t0.*

Initially, purification of the protein was an arduous task. Major obstacles were proteolytic degradation by enzymes in crude leucocyte extracts and loss due to binding of the protein to various surfaces in the presence of calcium. The protein was highly protease resistant in the presence of calcium; however, addition of calcium precluded the use of gel filtration columns or standard membranes to concentrate the dilute protein solutions by ultrafiltration. The protein could only be purified in stable form and in amounts sufficient for characterisation when PMSF was used to inhibit protease.[2] It was then possible to isolate and characterise the subunits.[5–7] Calprotectin (L1) has a molecular mass of about 36 kDa, and consists of subunit polypeptide chains, $L1_{Heavy}$ and $L1_{Light}$ (fig 2), with molecular masses of about 13 and 8 kDa, respectively. The N-terminal of the former is blocked, while the latter tends to form aggregates and precipitate when isolated.[7] The amino acid sequences of $L1_H$ and $L1_L$ are identical with the MRP-14 and MRP-8 proteins described by Odink *et al.*[8] The sequence of $L1_L$ is also identical with that of the cystic fibrosis antigen (CFAg).[9–11] Wilkinson *et al*,[11] using SDS electrophoresis, showed that CFAg was present on a protein containing subunits of 11 and 14 kDa. Although these authors mentioned that data from blot analysis under non-reducing conditions and gel filtration experiments suggested that the subunits occur predominantly as a heterodimer, they concluded that they are distinct proteins, and were named calgranulin A and calgranulin B. Subsequent reports by two other groups[12 13] have confirmed that the "cystic fibrosis protein" occurs as a complex, also known as p8,14 . In fact, no data have been reported as yet to suggest that either of the chains ($L1_H$ = MRP-14 = p14 = calgranulin B or $L1_L$ = MRP-8 = p8 = CFAg = calgranulin A) occur as free, individual proteins in any biological material. Gel filtration fractions of human plasma covering the molecular mass range from 0·1 to 600 kDa were tested by both enzyme linked immunosorbent assay and dot blot using polyclonal rabbit calprotectin antibodies directed against three different antigenic epitopes on each of the chains.[14] No reactivity was found in the fractions corresponding to a molecular mass below 36 kDa. The same was found when crude granulocyte extracts or purified calprotectin was tested. When the gel filtration buffer contains EDTA, about 85% of the calprotectin in plasma elutes at a molecular mass of about 36 kDa, and the rest between 100 and 300 kDa. If the buffer contains calcium, the major calprotectin peak occurs at 70 kDa, suggesting that the protein has dimerised (Fagerhol, unpublished data).

Calprotectin is also excreted in stools.[15] The median concentration in samples from healthy adults is about 3 mg/l. Free calprotectin subunits were not found (Røseth and Fagerhol, unpublished data).

On the basis of the data available, published and unpublished, it is reasonable to conclude that calprotectin is present as an oligomer or is complexed with other macromolecules. Free light and heavy chains can be seen only when samples are subjected to dissociating conditions—for example, during preparation for SDS electrophoresis.

What is the quaternary structure of calprotectin?

The experiments carried out by Edgeworth *et al*[12] suggested that p8,14 (calprotectin) has a 1:1 stoichiometry—that is, it is a heterodimer, although the presence of more than two subunits fits better with its estimated molecular mass of 35 kDa. Even if the subunit masses are taken to be 11 and 14 kDa, the highest values reported, the sum of a heterodimer seems too small. Based upon the masses of about 14 and 8 kDa, one light and two heavy chains gives exactly 36 kDa. A heterotrimer structure is also compatible with the presence of two calcium binding sites on each subunit[8 16] and the binding of up to six calcium ions in equilibrium dialysis experiments.[16A] Possible explanations for the discrepant conclusions may be found in how the cells were handled. Firstly, citrate, in contrast to EDTA, cannot prevent leakage of large

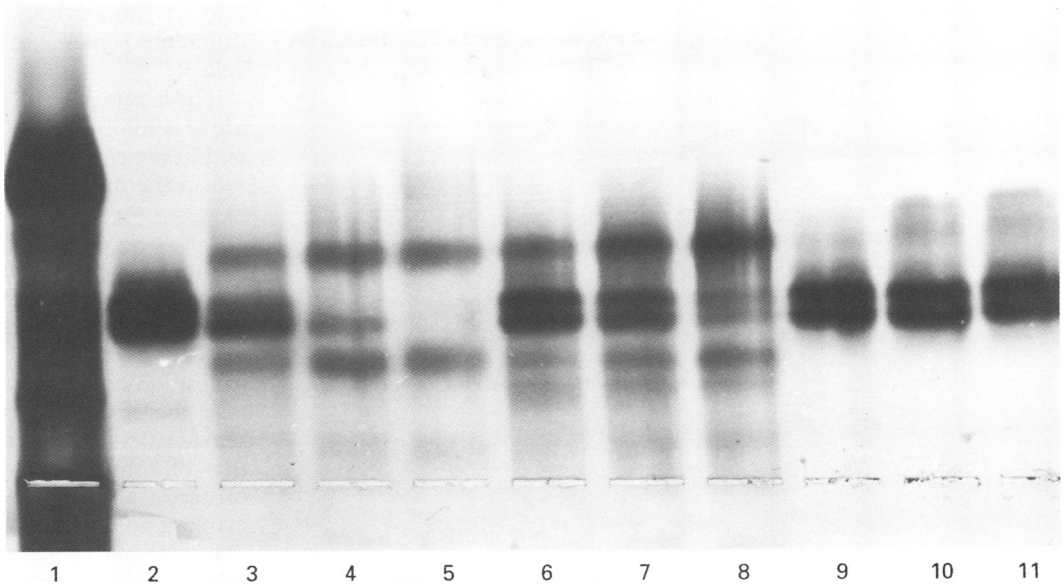

Figure 4 Banding pattern produced on agarose gel elctrophoresis of calprotectin incubated with lysosmal enzymes (E) from human granulocytes and inhibition of proteolysis by 10 mM PMSF. Samples: 1 = human serum; 2 = purified calprotectin (5 mg/ml); 3–11 = calprotectin incubated at 37°C for 30, 60 or 90 minutes in the presence of 15 mM sodium azide (samples 3–5), 0·25 mM thimerosal (samples 6–8) or 10 mM PMSF (samples 9–11).*

amounts of proteins from leucocytes when used as an anticoagulant.[3] Secondly, 1 mM PMSF may not be sufficient to inhibit protease (figs 3 and 4). If the extraction procedure involves incubation and contact of blood/cells with different types of materials and surfaces, leucocytes may become activated, rendering the protein unstable. This might explain loss of about 75% of the protein after two purification steps.[12] Protein loss from the cells may also explain the low estimate for the calprotectin content of monocytes. Such problems can be avoided by using EDTA as an anticoagulant or by adding EDTA (to a final concentration of 2·5 mM) to citrated blood within one hour of venepuncture. Neutrophil granulocytes can be rapidly isolated from blood by haemolysis with ammonium chloride and differential centrifugation.[1] Adequate protease inhibition is secured by addition of 10 mM PMSF[2] or preferably 10 mM DFP to the cell suspension buffer before neutrophil lysis. Large amounts of calprotectin can be purified from lysates of such cells by ion exchange chromatography on DEAE sepharose taking advantage of the appparent shift in the isoelectric point of the protein in the presence of EDTA (fig 5). When eluted from the chromatography column following the addition of calcium, calprotectin will be in its protease resistant conformation (fig 6).

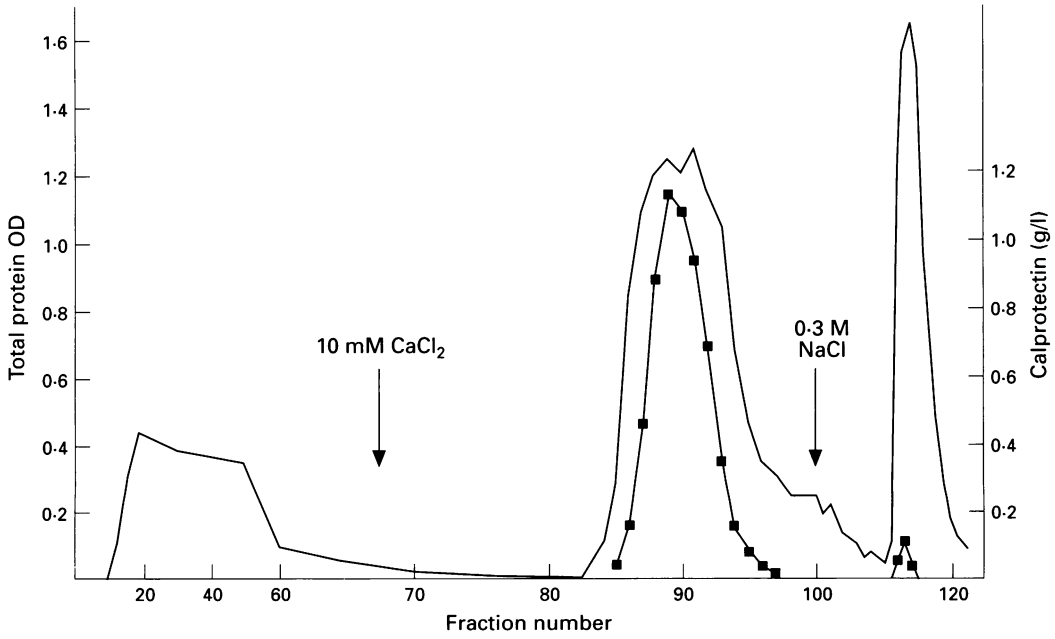

Figure 5 Purification of calprotectin from a crude, granulocyte lysate by ion exchange chromatography on a 2·5 × 10 cm column packed with DEAE-Sepharose FastFlow (Pharmacia, Sweden) equilibrated with 20 mM sodium barbital (pH 7·5) with 0·75 mM EDTA. Calprotectin (■) was eluted with 20 mM sodium barbital (pH 7·5) with 10 mM calcium chloride.

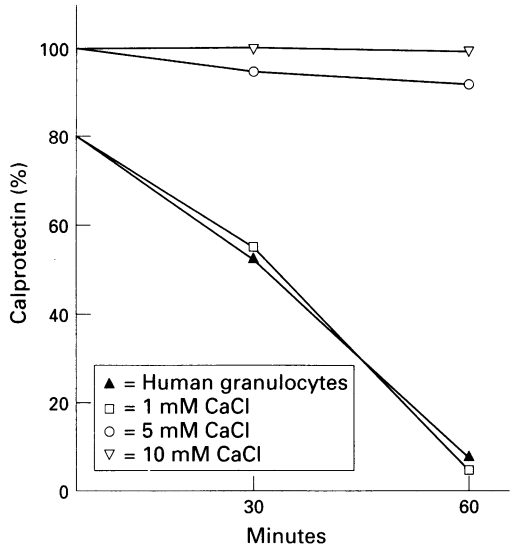

Figure 6 Degradation of calprotectin by lysosmal enzymes (E) from human granulocytes and the inhibitory effect of 1, 5 or 10 mM calcium chloride when incubated at 37°C for 30 to 60 minutes. Calprotectin concentrations are expressed as the per cent value at time t0.*

Is there a need for names other than calprotectin?

Separate names for calprotectin subunits should be considered if they exist independently and not just as laboratory artefacts, and if important biological functions are associated with the free light and heavy chains and not with calprotectin per se.

Calprotectin is the single descriptive name suggested for what most likely is a heterotrimeric protein. The name calprotectin reflects its calcium binding properties and its antimicrobial, anti-nematodal and thereby putative protective function.[4 13 17-19] The finding of very high concentrations of calprotectin in stools from patients with inflammatory bowel disease[15] suggests that it is brought into the gut lumen by migration of neutrophil granulocytes attracted by bacterial chemotactic substances. Thus, calprotectin may contribute to the regulation of the microbial flora in the gut. Other biological activities, mainly related to activation, migration and differentiation of myelomonocytic cells, have also been reported and have been reviewed recently.[20] However, much work had still to be done before the major biological functions of this protein can be specified. Inherited deficiency of the protein might be very telling in this respect, but no such case has been found as yet. None the less, it is most likely that the biological activities and distribution of this protein are protective in function.

The names calgranulin A and B were introduced for the subunits[11] with the intention of reducing the confusion surrounding the nomenclature of calcprotectin; at the time the names L1 light and heavy chains, MRP-8, MRP-14 and CFAg were used by others. Although the name calgranulin may be indicative of the presence of the protein in granulocytes (in the cytosol rather than in granules), it does not reflect its much wider distribution, namely also in monocytes, macrophages and many different types of squamous epithelial cells.[11 21-30]

The letters A and B in addition to calgranulin are less informative than the alternatives light and heavy chains or numerals 8 and 14, which reflect the molecular masses of the subunits.

Conclusions and suggestions

Nomenclature of proteins has often caused controversy for several obvious reasons. It would be very useful if the World Health Organisation or another international body could take the responsibility for working out consensus rules for protein nomenclature including that of complex oligomeric proteins like calprotectin and the rapidly growing number of cytokines.

In naming proteins, scientists should carefully consider data collected by different analytical methods, recognising their limitations, before suggesting names for putative new proteins. Preferably, names should be descriptive and relate to the proteins as they are found in what may be regarded as the native, biologically active form in cells, tissues or biological fluids.

In early stages when data are lacking, preliminary and brief symbols are useful, and should preferably reflect the origin and one or more properties of the polypeptide/protein. An example might be MRP-8, short for a myelomonocyte related protein with a molecular mass of 8 kDa. As several proteins may have similar masses, p8 alone may be ambiguous.

When further analyses have revealed that the protein is a subunit of the functional protein, the latter should be given a proper name. Using calprotectin as an example, it might be referred to as a hetero-oligomer consisting of the calprotectin p8 and p14 subunits.

1 Eggleton P, Gargan R, Fisher D. Rapid method for the isolation of neutrophils in high yield without the use of dextran or density gradient polymers. *J Immunol Methods* 1989;**212**:105–13.
2 Dale I, Fagerhol MK, Naesgaard I. Purification and partial characterization of a highly immunogenic human leukocyte protein, the L1 antigen. *Eur J Biochem* 1983;**134**: 1–6.
3 Fagerhol MK, Dale I, Andersson T. Release and quantitation of a leucocyte derived protein (L1). *Scand J Haematol* 1980;**24**:393–8.
4 Steinbakk M, Naess-Andresen CF, Lingaas E, Dale I, Brandtzaeg P, Fagerhol MK. Antimicrobial actions of calcium binding leucocyte L1 protein, calprotectin. *Lancet* 1990;**336**:763–7.
5 Andersson KB, Sletten K, Berntzen HB, Dale I, Brandtzaeg P, Jellum E, *et al.* The leucocyte L1 protein: identity with the cystic fibrosis antigen and the calcium-binding MRP-8 and MRP-14 macrophage components. *Scand J Immunol* 1988;**28**:241–5.
6 Berntzen HB, Fagerhol MK. L1, a major granulocytic protein: antigenic properties of its subunits. *Scand J Clin Lab Invest* 1988;**48**:647–52.
7 Berntzen HB, Fagerhol MK. L1, a major granulocyte protein; isolation of high quantities of its subunits. *Scand J Clin Lab Invest* 1990;**50**:769–74.
8 Odink K, Cerletti N, Brüggen J, Clerc RG, Tarcsay L, Zwadlo G, *et al.* Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis. *Nature* 1987;**330**: 80–2.
9 Wilson GB, Jahn TL, Fonseca JR. Demonstration of serum protein differences in cystic fibrosis by isoelectric focussing in thin-layer polyacrylamide gels. *Clin Chim Acta* 1973; **49**:79–85.
10 Van Heyningen V, Hayward C, Fletcher J, McAuley C. Tissue localisation and chromosomal assignment of a serum protein that tracks the cystic fibrosis gene. *Nature* 1985;**315**:513–15.
11 Wilkinson MM, Bussuttil A, Hayward C, Brock DJH, Dorin

JR, Van Heyningen V. Expression pattern of two related cystic fibrosis-associated calcium-binding proteins in normal and abnormal tissues. *J Cell Sci* 1988;**91**:221–30.
12 Edgeworth J, Gorman M, Bennett R, Freemont P, Hogg N. Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J Biol Chem* 1991;**266**:7706–13.
13 Barthe C, Figarella C, Carrère J, Guy-Crotte O. Identification of "cystic fibrosis protein" as a complex of two calcium-binding proteins present in human cells of myeloid origin. *Biochim Biophys Acta* 1991;**1096**:175–7.
14 Berntzen HB, Endresen GKM, Fagerhol MK, Spiechowitz J, Mowinckel P. Calprotectin (the L1 protein) during surgery in patients with rheumatoid arthritis. *Scand J Clin Lab Invest* 1991;**51**:643–50.
15 Røseth AG, Fagerhol MK, Aadland E, Schjønsby H. Assessment of the neutrophil dominating protein calprotectin in feces. A methodological study. *Scand J Gastroenterol* 1992;**27**:793–8.
16 Lagasse E, Clerc RG. Cloning and expression of two human genes encoding calcium-binding proteins that are regulated during myeloid differentiation. *Mol Cell Biol* 1988; **8**:2402–10.
16A Naess-Andresen CF, Egelandsdal B, Fagerhol MK. Calcium binding and concomitant changes in the structure and heat stability of calprotectin (L1-protein). *J Clin Pathol: Mol Pathol* 1995;**48**:M278–84.
17 Sohnle PG, Collins-Lech C, Wiessner JH. Antimicrobial activity of an abundant calcium-binding protein in the cytoplasm of human neutrophils. *J Infect Dis* 1991;**163**: 187–92.
18 Murthy ARK, Lehrer RI, Harwig SSL, Miyasaki KT. In vitro candidastatic properties of the human neutrophil calprotectin complex. *J Immunol* 1993;**151**:6291–301.
19 Edgeworth JD, Abiose A, Jones BR. An immunohistochemical analysis of onchocercal nodules: evidence for an interaction between macrophage MRP8/MRP14 and adult Onchocerca volvulus. *Clin Exp Immunol* 1993;**92**:84–92.
20 Hessian PA, Edgeworth J, Hogg N. MRP-8 and MRP-14, two abundant Ca^{2+}-binding proteins of neutrophils and monocytes. *J Leukoc Biol* 1993;**53**:197–204.
21 Gabrielsen TØ, Dale I, Brandtzaeg P, Hoel PS, Fagerhol MK, Larsen TE, *et al*. Epidermal and dermal distribution of a myelomonocytic antigen (L1) shared by epithelial cells in various inflammatory skin diseases. *J Am Acad Dermatol* 1986;**15**:173–9.
22 Brandtzaeg P, Dale I, Fagerhol MK. Distribution of formalin-resistant myelomonocytic antigen (L1) in human tissues. II. Normal and aberrant occurence in various epithelia. *Am J Clin Pathol* 1987;**87**:700–7.
23 Gabrielsen TØ, Brandtzaeg P, Hoel PS, Dale I. Epithelial distribution of a myelomonocytic antigen L1 in relation to cutaneous malignancies and melanocytic naevi. *Br J Dermatol* 1988;**118**:59–67.
24 Kelly SE, Jones DB, Fleming S. Calgranulin expression in inflammatory dermatoses. *J Pathol* 1989;**159**:17–21.
25 Dale I, Brandtzaeg P. Expression of epithelial L1 antigen as an immunohistochemical marker of squamous cell carcinomas of the lung. *Histopathology* 1989;**14**:493–502.
26 Clark BR, Kelly SE, Fleming S. Calgranulin expression and association with the keratinocyte cytoskeleton. *J Pathol* 1990;**160**:25–30.
27 Tungekar MF, Heryet A, Gatter KC. The L1 antigen and squamous metaplasia in the bladder. *Histopathology* 1991; **19**:245–50.
28 Eversole LR, Miyasaki KT, Christensen RE. The distribution of the antimicrobial protein, calprotectin, in normal oral keratinocytes. *Arch Oral Biol* 1992;**37**:963–8.
29 Kunz M, Roth J, Sorg C, Kolde G. Epidermal expression of the calcium binding surface antigen 27E10 in inflammatory skin diseases. *Arch Dermatol Res* 1992;**284**: 386–90.
30 Miyasaki KT, Bodeau AL, Murthy ARK, Lehrer RI. In vitro antimicrobial activity of the human neutrophil cytosolic S-100 protein complex, calprotectin, against capnocytophaga sputigena. *J Dent Res* 1993;**72**:517–23.