

Primary structure of bovine lactoperoxidase, a fourth member of a mammalian heme peroxidase family

Marie-Madeleine CALS, Pascal MAILLIART, Ghislaine BRIGNON, Patricia ANGLADE and Bruno Ribadeau DUMAS
Protein Unit, Institut National de la Recherche Agronomique, Jouy-en-Josas, France

(Received September 3, 1990/February 19, 1991) — EJB 90 1047

Much is known about bovine lactoperoxidase but no data are available on its primary structure. In this work its main active fraction was isolated from cow's milk and sequenced using a conventional strategy. A clear similarity was found with human myeloperoxidase, eosinophil peroxidase and thyroperoxidase, the sequences of which were recently elucidated from those of their cDNAs and/or genes. The single peptide chain of bovine lactoperoxidase contains 612 amino acid residues, including 15 half-cystines and 4 or 5 potential *N*-glycosylation sites. The corresponding peptide segments of human myeloperoxidase, eosinophil peroxidase and thyroperoxidase display 55%, 54% and 45% identity with bovine lactoperoxidase, respectively, with 14 out of the 15 half-cystines present in each of the four enzymes being located in identical positions. The occurrence of an odd number of half-cystines in bovine lactoperoxidase supports the recent finding of a heme thiol released from this enzyme by a reducing agent, suggesting that the heme is bound to the peptide chain via a disulfide linkage, since the absence of free thiol in the enzyme was reported long ago.

Closely related heme peroxidases have been detected in most mammalian exocrine secretions. They are believed to protect mucosal surfaces from microorganisms by catalyzing the production of toxic oxidizing agents from H_2O_2 , and halides or SCN^- , a pseudohalide. Two of them, bovine lactoperoxidase and human salivary peroxidase, have been especially studied (Tenovuo, 1985). Bovine lactoperoxidase was highly purified and characterized in the early forties (Theorell and Akeson, 1943). Although lactoperoxidase is able to oxidize Br^- , I^- and SCN^- , Reiter and coworkers (1964) demonstrated that oxidation of SCN^- was involved in the mechanism of the lactoperoxidase-catalyzed antibacterial effect in bovine milk. Aune and Thomas (1977) and Hoogenboom et al. (1977) independently concluded that hypothiocyanite (OSCN^-) was the main antibacterial species produced from SCN^- and H_2O_2 .

Bovine lactoperoxidase consists of a single peptide chain (Sievers, 1981), of molecular mass 70 kDa, a heme and about 10% carbohydrate (Rombauts et al., 1967; Carlström, 1969). It was recently shown that the heme was protohemin IX with a mercaptomethyl group replacing a methyl group. This suggested that the protoporphyrin was linked to the peptide chain through a disulfide bridge with a cysteine residue

(Nichol et al., 1987). Recently a calcium ion was found to be strongly bound to bovine lactoperoxidase (Booth et al., 1989). The electrophoretic heterogeneity of bovine lactoperoxidase was carefully studied by Carlström (1969). No significant difference was found between the amino acid compositions of its various active fractions. However, some of them displayed lower carbohydrate content and molecular mass. Partial deamidation and/or proteolysis were suggested to contribute to heterogeneity which was found even in lactoperoxidase from individual cows.

We report here the primary structure determination of the main bovine lactoperoxidase fraction and show the striking similarity this enzyme displays with human leucocyte peroxidases (myeloperoxidase and eosinophil peroxidase) and human thyroperoxidase that have been recently sequenced from cDNAs and/or genes (myeloperoxidase: Morishita et al., 1987a, b; Johnson et al., 1987, 1989; eosinophil peroxidase: Ten et al., 1989; Sakamaki et al., 1989; thyroperoxidase: Libert et al., 1987; Kimura et al., 1987, 1989). Porcine (Magnusson et al., 1987) and rat (Derwahl et al., 1989; Isosaki et al., 1989) thyroperoxidases as well as murine myeloperoxidase (Venturelli et al., 1989a, b) have also been similarly sequenced recently.

Correspondence to B. Ribadeau Dumas, Protein Unit, Bat. 224, INRA, F-78352 Jouy-en-Josas Cedex, France

Abbreviations. Ptc, phenylthiocarbamoyl; Pth, phenylthiohydantoin; Pe, pyridylethyl; Cm, *S*-carboxymethyl; RP, reverse phase.

Enzymes. Lactoperoxidase, myeloperoxidase, eosinophil peroxidase, thyroperoxidase (EC 1.11.1.7); endoproteinase Lys-C (EC 3.4.21.50); endoproteinase Glu-C (EC 3.4.21.19); trypsin (EC 3.4.21.4); carboxypeptidase P (EC 3.4.16.1).

Note. The novel amino acid sequence data published here have been deposited with the EMBL sequence data bank and is available under accession number P 80025.

MATERIALS AND METHODS

Assays for peroxidase activity were carried out according to Bardsley (1985), with diammonium 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonate) as substrate. The activity unit is the amount of enzyme which catalyzes the oxidation of 1 μmol substrate/min at pH 5.5, 25°C.

The procedure used to prepare bovine lactoperoxidase (Mailliart, 1989) will be published elsewhere and is briefly described here. The starting material was fresh, defatted bulk

rennet-lactoserum, the proteins of which were 10-fold concentrated by ultrafiltration. These proteins were adsorbed at neutral pH on a carboxymethyl-Sepharose column and eluted stepwise with two 80 mM ammonium acetate solutions adjusted with ammonia to pH 8.5 and 10.5, respectively. After thorough washing with the former, bovine lactoperoxidase was eluted by the latter and directly freeze-dried. It was further fractionated in two steps by FPLC (Pharmacia). The first step (Mono Q column; 25 mM triethylamine pH 10.5; 0–0.6 M NaCl gradient) gave three active fractions: A (the main one), B and C, in order of elution. In the second step, A was chromatographed on a Mono S column (50 mM ammonium acetate pH 8.0; 0.2–0.3 M NaCl gradient), giving two active sub-fractions, A1 (the main one) and A2, in order of elution. The complete sequence was obtained from A1, while only the N-terminus of A2, as well as that of bovine lactoperoxidase from Sigma, were sequenced.

Amino acid analyses were performed on A1, A2 and some peptides after acid hydrolysis (liquid or gaseous 5.7 M HCl, under vacuum or argon; 110°C; hydrolysis times: 24, 48, 96 h for proteins, 24 h for peptides), either with a Biotronik LC 5000 analyzer (Munich, FRG), or, according to Bidlingmeyer et al. (1984), after derivatization with phenylisothiocyanate. Tryptophan was determined according to Yokote et al. (1988), with hydrolysis times of 25, 50 and 75 min at 166°C.

A1 was reduced with 2-mercaptoethanol and alkylated by iodoacetate or 4-vinyl pyridine (Henschen, 1986) giving *S*-carboxymethyl A1 (Cm-A1) or pyridylethyl A1 (Pe-A1), respectively. Several digestions of the alkylated molecule were performed: using endoproteinase Lys-C (Boehringer; 25 mM Tris, 1 mM EDTA, 4 M urea, pH 8.5; 37°C; enzyme/substrate, E/S = 1:500 by mass; enzyme added in three aliquots at $t = 0$ h, 1 h, 2 h; total digestion time: 16 h); endoproteinase Glu-C (Boehringer; 50 mM ammonium acetate, 2 mM EDTA, pH 4.0; 37°C; E/S = 1:30 by mass; 5 h); trypsin (treated with tosylphenylalaninechloromethane) after blocking amino groups with phenylisothiocyanate (Tarr, 1982) (0.2 M ammonium bicarbonate, pH 7.8; 37°C; E/S = 1:100 by mass; overnight); CNBr (70% trifluoroacetic acid; 0.15 mg CNBr/ml; molar ratio CNBr/Met = 100; room temperature; overnight); and hydroxylamine (Bornstein and Balian, 1977). Several purified peptides were digested with trypsin; their fragments were purified and analyzed. Alkylated A1 and some peptides were treated with carboxypeptidase P (Lu et al., 1988). Amino acid analyses were performed on the reaction mixtures after 6 h.

The various digests were fractionated by RP-HPLC using Waters equipment (see Supplement at the end of the paper). Two types of columns were used: Aquapore RP-300 (Brownlee; 30 × 4.6 mm, C₄ or C₈, 7 µm, 30 nm) and Nucleosil (SFCC, FR; 250 × 4.6 mm, C₁₈, 5 µm, 10 nm). The two following elution systems were employed: (I) solution A = 0.115% trifluoroacetic acid; solution B = 0.1% trifluoroacetic acid, 80% CH₃CN or 80% CH₃CN/2-propanol (1:1, by vol.); (II) solution A = 25 mM ammonium acetate pH 6.0; solution B = 50 mM ammonium acetate pH 6.0/80% CH₃CN/2-propanol (1:1, by vol.). Linear gradients (single or multiple slopes) from A to B were used for the elution at 1 ml/min, with absorbance monitored at 214 nm. On many occasions peptides were purified using successively two different solvent systems. Collected fractions were dried using a Speed Vac concentrator (Savant).

Sequence determinations were performed on a 477 A sequencer coupled to a 120 A HPLC (Applied Biosystems), using solvents supplied by the manufacturer.

Table 1. *Amino acid composition of bovine lactoperoxidase*

All figures are rounded to the nearest integer. The values for LPO-B1 (obtained by Carlstrom, 1969) were for hydrolyses in duplicate after 20 h, 48 h, 72 h and 140 h, using a molecular mass of 78 500 Da; titration showed the absence of free cysteine. Values for LP-A1 and LP-A2 were obtained in this work. The first column of values for LP-A1 were obtained by sequencing by Edman degradation. Calculated values for LP-A1 assumed 612 residues; calculated values for LP-A2 assumed 596 residues, taking into account a 16-residue N-terminal deletion (Asx₄, Thr⁵, Ser, Pro, Ala, Val, Leu₃ from the sequence). the resulting molecular mass of the peptide chain calculated from the sequence in this work was 69 502 Da; that calculated by Carlstrom (1969), which included carbohydrates and heme, was 78 030 Da

Amino acid	Residues in			
	LPO-B1	sequence LP-A1	LP-A1	LP-A2
Asp		35		
Asx	71	71	72	67
Asn		36		
Thr	32	32	32	27
Ser	33	32	34	34
Glu		33		
Glx	61	59	62	63
Gln		27		
Pro	42	42	43	43
Gly	40	40	40	40
Ala	37	38	38	38
1/2 Cys	16 ^c	15	11	8
Val	28	29	30	29
Met	11	11	11	11
Ile	27	27	24	25
Leu	68	69	67	63
Tyr	15	15	15	15
Phe	30	31	30	31
Lys	34	33	35	35
His	14	14	16	15
Arg	37	39	38	38
Trp	15	14	13	13
Total	611	613	611	595
Man	26	—	—	—
GlcNAc	14	—	—	—
GalNAc	4	—	—	—

RESULTS

Part of this work was included in a thesis (Maillart, 1989), notably the procedure used to purify and fractionate bovine lactoperoxidase, and some sequence data. The latter were obtained by directly sequencing Cm-A1 and A2; by sequencing peptides isolated from a CNBr digest of Cm-A1 and a tryptic digest of phenylthiocarbamoylated Cm-A1 (Ptc-Cm-A1). As mistakes were found later, concerning in particular Gln and Cys since Pth-Gln and Pth-Cm-Cys often coeluted, the data mentioned above were reinvestigated and completed, using endoproteinase Glu-C, CNBr and hydroxylamine digests of Pe-A1 as well as an endoproteinase Lys-C digest of Cm-A1.

The mobilities of the single bands obtained with A1 and A2 by SDS/PAGE were identical (data not shown). For these fractions the ratios $\epsilon_{412}/\epsilon_{280}$ were 0.89 and 0.93 and the specific activities 227 and 234 U/mg, respectively. A2 obviously contained several forms of lactoperoxidase, as indicated by the elution pattern from the Mono S column and the electrophoretic pattern [PAGE, pH 4.3, staining with

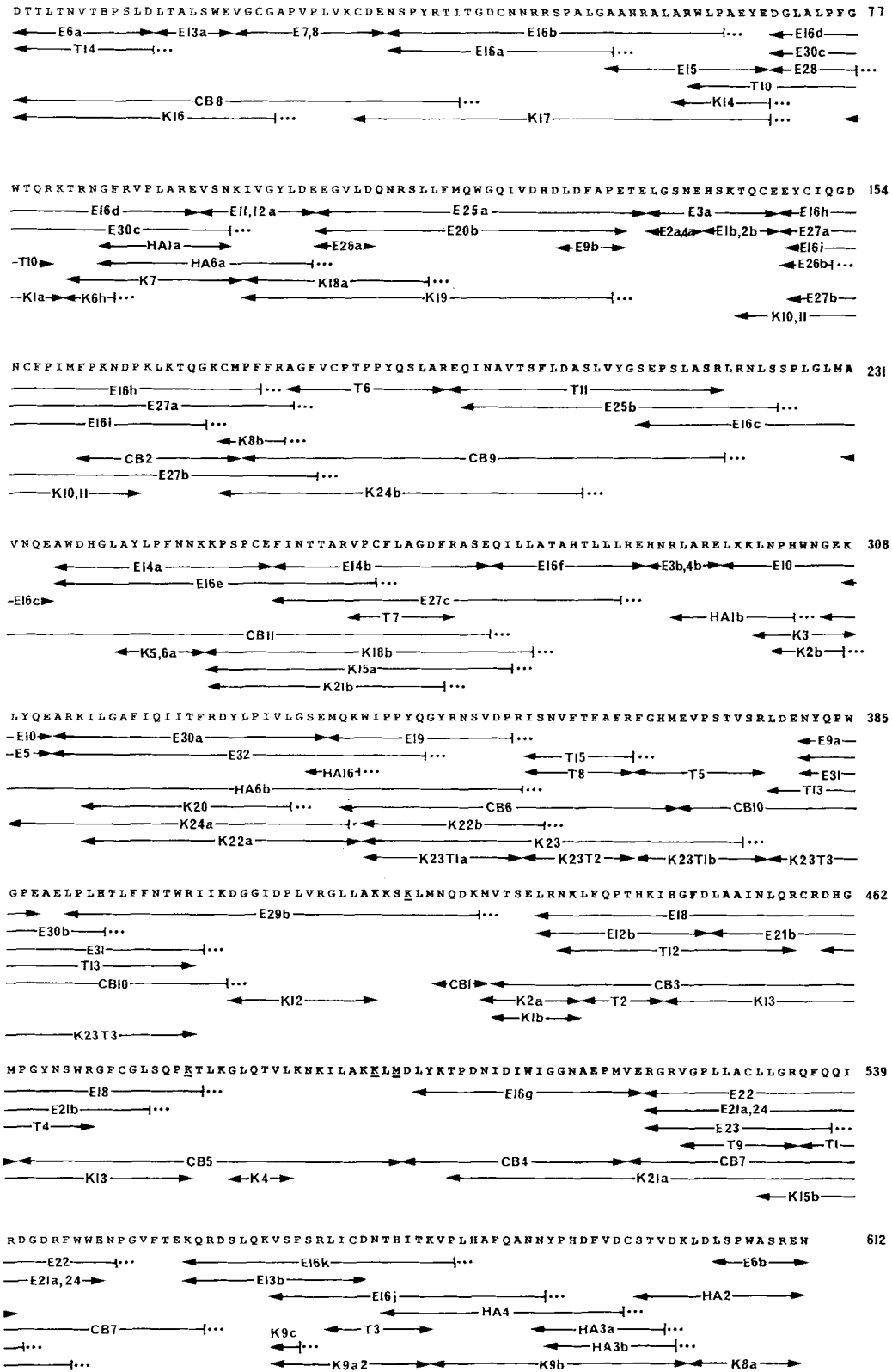


Fig. 1. Alignment and overlapping of the peptides obtained from the different hydrolysates. K, E, CB, HA and T, peptides obtained from the endoproteinase Lys-C, endoproteinase Glu-C, CNBr, hydroxylamine and tryptic digests (of Ptc-Cm-A1 for the latter), respectively. Peptide numbering (first figure after one of the letters shown above) corresponds to the order of elution in the first fractionation of a digest. The four underlined residues were not identified with certainty

diammonium 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonate) and H_2O_2]. Two main fractions and two minor ones were visible using this procedure. The latter had the lowest mobility while the most abundant fraction had the highest (data not shown).

The amino acid compositions of A1 and A2 are shown in Table 1. The first 27 residues of Cm-A1 were directly sequenced. The first residue was not clearly identified (probably Asp). No Pth-Xaa was detected at position 6. Residue 23 could be either Gln or Cm-Cys. Fourteen Edman cycles were performed on A2. Two sequences were found, starting at positions 16 and 18 of A1. Bovine lactoperoxidase from Sigma gave the same two sequences. Carboxypeptidase P released Asn, Ser, Glu and Arg from Pe-A1.

Fig. 1 shows how the sequence of A1 was deduced from the study of the various hydrolysates, also taking into account the N- and C-terminal determinations carried out on the entire protein. Fig. 2 compares the 612-residue peptide chain of bovine lactoperoxidase with the corresponding regions of human myeloperoxidase, eosinophil peroxidase and thyroperoxidase. For a given peptide, some residues could not be identified with certainty when only sequence data obtained from this peptide were taken into account. Obviously the number of such residues progressively diminished when results originating from other peptides or other digests were combined. In Fig. 1, four of them remain (positions 422, 479, 496, 498). In addition, peptides overlapping the two bonds 135–136 and 289–290 were not found. The similarities between bovine lactoperoxidase and the three human enzymes (myeloperoxidase, eosinophil peroxidase, thyroperoxidase) amount to 55.4%, 54.0% and 44.6%, respectively. Of the 15 half-cystines, 14 are in identical positions in the human peroxidases. The 15th is located at position 458. Four or five putative N-glycosylation sites are located at positions 6, 9, 112, 222 and 258. According to the peptides studied, the following Asx residues were found during sequencing: position 6, Asn, minor Asp or nothing; position 9, Asp or Asn; positions 112, 222, 258, Asn or nothing.

DISCUSSION

The isolation, molecular mass, heterogeneity, amino acid composition, enzymatic properties and biological function of bovine lactoperoxidase have been extensively studied. Its high isoelectric point (> 8) makes its separation from the main lactosoluble proteins easy by cation exchange. However contaminants of similar isoelectric points (immunoglobulins, lactoferrin, lysozyme) must be discarded by further purifications. More than ten active fractions were isolated and characterized by amino acid, carbohydrate, heme and molecular mass determinations (Carlström, 1969). Among them, some clearly had different carbohydrate contents. Partial deamidation and/or proteolysis were suggested to participate in bovine lactoperoxidase heterogeneity, while the contribution of genetic variations was discarded by studies carried on milk from single animals.

We have isolated the main active fraction, A1, which appeared to be homogeneous according to several criteria, and another, A2, devoid of inactive contaminant proteins but containing two largely predominant active species, as judged by SDS/PAGE, and PAGE at pH 4.3 with staining for activity. A1 is probably identical to the fraction LPO-B1 studied by Carlström (1969). Its amino acid composition is close to that determined on this fraction by Carlström. However we think

that the values he obtained for 1/2Cys, Val and Ile are likely to be more accurate than ours. Indeed Carlström determined the former after periodate oxidation, whereas the figures we give were obtained as cystine on the native protein, which is certainly not the best way. On the other hand, the values for Val and Ile were obtained after hydrolysis times of 140 h by Carlström and 96 h by us. As far as A2 is concerned, a slightly shorter peptide chain could be expected.

Bovine lactoperoxidase is known to have a single peptide chain (Sievers, 1981) with about 10% carbohydrates (Carlström, 1969). An N-terminal leucine was found by Sievers (1980). Most ambiguities present in the N-terminal sequence of A1 obtained directly were eliminated by the study of peptides from different hydrolysates: the N-terminus is Asp; cysteine occurs at position 23; residue 6 is Asn, probably glycosylated. However, in our final sequence (Fig. 1), B at position 9 stands for Asn or Asp: indeed, we found either of them in high yield at this position, according to the peptide or protein under study, but never a blank which could signal the presence of glycosylated Asn. The presence of Asp could be due either to deamidation of Asn, or to the occurrence of two alleles (Asp $\leftrightarrow$ Asn), resulting from a point mutation, in our preparation. In any case Asn9 is not glycosylated. C-terminal peptides were found in three digests (Fig. 1). They were the only peptides which liberated the same C-terminal amino acids as Pe-A1 through carboxypeptidase P action.

Most of the overlaps between the various peptides we have studied are satisfactory (Fig. 1). Only one Met makes the overlap at position 429. But, taking into account the almost certain occurrence of Met at position 498, since a CNBr peptide starts at 499, and the identical number of Met residues found from the sequence and amino acid analysis (Table 1), we conclude that the positioning of peptides in this area is correct. Since only two missing overlaps are left, there is only one way of ordering the three corresponding peptide segments, taking into account the knowledge of the positioning of two of them (N- and C-terminal segments) from the study of both ends of Cm-A1 (see above). Thus there is no need to take advantage of similarities with other peroxidases to present the sequence of bovine lactoperoxidase as shown in Fig. 1. Although peptide segments located between residues 135–136 and/or 289–290 may have escaped our attention, this is highly unlikely: indeed a direct linkage in each case fits endoproteinase Glu-C specificity. Furthermore, the amino acid composition of the whole protein (Table 1) is almost identical with that deduced from the sequence. We believe that the only weakness of our results lies in the uncertain identification of the four residues mentioned in the Results. Both Ala and Ser were found at position 294, probably arising from a point mutation.

As far as A1 is concerned, the molecule has 612 residues. It can be aligned with human myeloperoxidase, eosinophil peroxidase and thyroperoxidase quite satisfactorily with very few insertions/deletions (Fig. 2). The similarity is highest with human myeloperoxidase and eosinophil peroxidase (55.4% and 54%, respectively) and lower (44.6%) with human thyroperoxidase. Clearly, lactoperoxidase is encoded by a gene different from those encoding myeloperoxidase, eosinophil peroxidase and thyroperoxidase since similarities between human and murine myeloperoxidases (Venturelli et al., 1989a), human, porcine (Magnusson, 1987) and rat thyroperoxidases (Derwahl et al., 1989; Isosaki et al. 1989) are much higher. If the peptide segments corresponding to the whole length of lactoperoxidase are considered, each peroxidase has 15 half-cystines, 14 being at the same position. The non-aligned half-

hTP	***TQQSQHP	TDAISEDLS	IIANMSGCLP	YMLPPKCPNT	CLANKYRPIT	GACNNRDHPR	165
hEP	***RSGGFNV	TDVLTEPQLR	LSQASGCAL	RQDAERCS--	---DKYRTIT	GRCNKKRPL	159
hMP	***WRRPFNV	TDVLTPAQLN	VLSKSSGCAY	QDVGVTCPEQ	---DKYRTIT	GMCNNRRSPIT	187
bLP	DTTLNV	TBP--SLDIT	ALSWEVGCGA	PVPLVKCDEN	SP--YRTIT	GDCNNRSPA	52
hTP	WGASNTALAR	WLPVYEDGF	SQPRGMNPGF	LYNGFPLPPV	REVTTRHVIQV	SNEVVTDDDR	225
hEP	LGASNOALAR	WLPAYEDGL	SLPFGWTPSR	RNRGFLFLPLV	RAVSNQIVRF	PNERLTSDRG	219
hMP	LGASNRAFVR	WLPAYEDGF	SLPYGWTGPGV	KRNGFPVALLA	RAVSNQIVRF	PTDOLTPDQE	247
bLP	LGASNRAFVR	WLPAYEDGL	ALPFGWTPSR	TRNGFRVLLA	REVSNTLVGY	LDEEGVLQDN	112
hTP	YSDLLMANGQ	YIDHDIATFR	OSTSKAAFGG	GADQMTCEEN	QNPFFPIQLP	EARPAAGTA	285
hEP	RATLMEQWQO	FIDHDLDFSH	ESPARVAFTA	GVDCEITCAQ	LPPCFPIKIP	PNDPRIRKQR	279
hMP	RSLMEQWQO	LIDHDLDFTP	EPARASFTV	GVNCEITSCVQ	QPPCFPLKIP	PNDPRIRKQA	307
bLP	RSLMEQWQO	IVDHDLDFAF	ETELGSNEHS	KTCCEYCIQ	GDNCFPIIMP	KNDPRIRKQ	172
hTP	CLPFYRSSA	ACGTGDOGAL	FGNLSTANPR	QMNGLTSFL	DASTVYGSSE	ALERQLRNMT	344
hEP	DCIPFFRSAP	SCFQNKNRV--	-----R	NOINALTSTFV	DASMYGSEV	SLSLRLRNMT	329
hMP	DCIPFFRSCP	ACPGSNITI--	-----R	NOINALTSTFV	DASMYGSEE	PLARNLRNMS	357
bLP	KCMFFRAGF	VCPTPPYQSL	A-----R	EOINAVTSFL	DASLYGSEP	SLASLRNLNS	224
hTP	SAEGLLRVHA	RLRDSGRAYL	FFVPPRRPAA	CAPEPGIPGE	TRGFCFLAGD	GRASEVPSLT	404
hEP	NYLGLLAINQ	RFQDMGRALL	PFDMHLDD--P	CLLTNRS--	ARIPCFLAGD	TRSTETPKLA	385
hMP	NQLGLLAINQ	RFQDMGRALL	PFDMHLDD--P	CLLTNRS--	ARIPCFLAGD	TRSEMPBLT	413
bLP	SFLGLLAINQ	EAWDHGLAYL	FFNNKKPS--P	CEFINNT--	ARVPCFLAGD	FRASECIILIA	280
hTP	ALHTMLREH	NRLAAAKKAT	NAHWSADAVY	QEARVVUGAL	HOIITLRYDI	PRILGPEAFQ	464
hEP	AMHTLMREH	NRLATELRRL	NPRWNGDRLY	NEARKIMGAM	VOIITYRDFP	PLVLGKARAR	445
hMP	SMHTLMLREH	NRLATELRKL	NPRWNGEGLY	QEARIVGAM	VOIITYRDYL	PLVLGPTAMR	473
bLP	TAHTLMLREH	NRLATELRKL	NPRWNGEGLY	QEARILGAF	VOIITYRDFP	PLVLGSEMQK	340
hTP	QVVGVEGMD	STAMPVSNV	FSTAFAFRGH	ATIHLVLRRL	DASFQEHFDL	PGLWHLQAF	524
hEP	RTLGHYRGYC	SMVDFRVANV	FTL-AFRFGH	TMLQPFMFRL	DSQFASASRN	SHVFISSAFF	504
hMP	KYLPTYSRYN	DSVDFRIANV	FTN-AFRFGH	TLIQPFMFRL	DNRYPQMEHN	PRVPLSRVFF	532
bLP	WIPPYQ-GYR	NSVDFRISNV	FTF-AFRFGH	MEVPSVSRRL	DENYQPWGHE	AELPLHTLFF	398
hTP	SPWLLRGGG	IDPLIRGLLA	RPAKLQVQDQ	LMNEELTERD	FVLSNSS--TL	DASINLQRG	583
hEP	ASWRIVYEGG	IDPILRGLMA	TPAKLNRODA	MLVDELDRRL	FRQVRRI--GL	DIAALNMQRS	563
hMP	ASWRVVEGG	IDPILRGLMA	TPAKLNRODA	IADDELIRRL	FEQVMRI--GL	DLPALNMQRS	591
bLP	NTWRIKDDG	IDPLVRGLLA	KKSILMMQDK	MVTSELIRNKL	EQPTHKTHGF	DIAATNLQRC	458
hTP	RDHGLFGYNE	WRFCGLPRL	ETPADLSTAI	ASRSVADKIL	DLYKHEDNID	VMIIGLAENF	643
hEP	RDHGLPGYNA	WRFCGLSP	RNLALSRVLI	KNODLARKFL	NLYGTFDNID	IWIIGIAEHL	623
hMP	RDHGLPGYNA	WRFCGLPQP	ETVGLGTVL	RNLKLARKLM	EQYGTFDNID	IWMGGVSEHL	651
bLP	RDHSGMEGYS	WRFCGLSP	KTLLKGLTVL	KMKILAKKLM	DLYMTFDNID	IWIIGNAEHL	518
hTP	LPRARTEGDF	ACLIGOMKA	LRDGRFWWE	NSHVFTEAQR	REDEKHSLSR	VICDNTGLTR	703
hEP	LPGARVGPLL	ACLIFENQRR	ARDGDRFWQ	KRGVTEKRR	KALLSRISLSR	IICDNTGITIT	683
hMP	KKGRVGPLL	ACIIGTQFRK	LRDGRFWWE	NEGVSQMR	QALLAQISLPR	IICDNTGITIT	711
bLP	VERGRVGPLL	ACLLGRQFQQ	LRDGRFWWE	NHGVTEKQR	DSLQKVSFPR	IICDNTHTITK	578
hTP	VPM-DAFQVG	KFFEDFESCD	SIPGMNLEAH	RETFP....			933
hEP	VSR-DIFRAN	IVPRGFVNC	RIPRNLISAW	RG			715
hMP	VSKNNIFMSN	SVPRDFVNC	TLPALNLSAW	REAS			745
bLP	VPL-HAFQAN	NYPHDFVDC	TVDKLIDUSPW	ASREN			612

Fig. 2. Comparison of the amino acid sequences of human thyroperoxidase (hTP, Kimura et al., 1987), human eosinophil peroxidase (hEP, Sakamaki et al., 1989), human myeloperoxidase (hMP, Morishita et al., 1987a) and bovine lactoperoxidase (bLP, this work). The first three sequences are numbered from Met1 while bLP is from the N-terminus of the longest species we have found. Identical residues in hTP, hEP, bLP, and in the four proteins are boxed. B, Asp or Asn; □, N-termini of the light and heavy chains of EP and MP; L, N-termini found in LP-A2 and in the Sigma enzyme; underlined segments, putative locations of the proximal and distal histidine regions (Kimura and Ikeda-Saito, 1988), ▲, LP putative N-glycosylation sites; underlined residues (positions 422, 479, 496, 498 of bLP), not identified with certainty

cystines in thyroperoxidase, eosinophil peroxidase, myeloperoxidase and lactoperoxidase occur at positions 36, 181, 349 and 458 (lactoperoxidase numbering as below), respectively. It was shown earlier (Carlström, 1969) that bovine lactoperoxidase had no free cysteine. Our results, showing the presence of an odd number of half-cystines, are in accordance with those of Nichol et al. (1987), who isolated from reduced bovine lactoperoxidase a protohemin IX thiol and suggested the occurrence of a disulfide bridge between the heme and the peptide chain. Indeed this was the first indication for the presence of a covalent linkage between the heme and the peptide chain.

No data are available concerning the binding of the heme in myeloperoxidase, eosinophil peroxidase and thyroperoxidase. It is believed that the fifth iron ligand in heme peroxidases is a histidine residue (proximal His), while another His plays a role in the catalysis (distal His). Their putative localization are in the sequences 272–294, and 365–368 or 457–461, respectively (Kimura and Ikeda-Saito, 1988). Indeed only four His residues are in identical positions in the four peroxidases, two in the former sequence (residues 283 and 290) and one in each of the latter two (residues 368 and

461). From comparison with non-mammalian heme peroxidases, Kimura and Ikeda-Saito (1988) suggested that His283 could be the proximal His. As far as the distal His is concerned, the two areas mentioned above are based on a comparison of yeast and plant peroxidases made by Poulos and Kraut (1980) who found the distal His located in a consensus sequence Arg-Xaa-Yaa-Zaa-His. Although sequence 457–461 exactly corresponds to this, Kimura and Ikeda-Saito (1988) hypothesized that the distal heme was at position 368 since they found a shorter human thyroperoxidase mRNA, which is produced as a result of alternate splicing and is still expressed. This mRNA lacks the consensus sequence 457–461 but contains sequence 365–368 which is close to the consensus. Our results do not allow us to settle the question.

As indicated in Results, four or five (see above) putative N-glycosylation sites occur in lactoperoxidase A1 (positions 6, 9, 112, 222 and 258). The first two sites are lacking in lactoperoxidase A2 which has Ala16 and Ser18 as N-termini. Our results seem to show that Asn residues 6, 112, 222 and 258 are capable of being glycosylated since, on several occasions, no Pth-Xaa could be detected at these positions. Asn222 is a putative glycosylation site in all the peroxidases

compared in Fig. 2, while sites 6, 112 and 258 are shared by bovine lactoperoxidase, human myeloperoxidase and human eosinophil peroxidase.

Recently Booth et al. (1989) found that both bovine lactoperoxidase and myeloperoxidase contained a tightly bound calcium ion. Computer amino acid sequence analyses and comparisons between human myeloperoxidase and calcium binding proteins revealed two plausible calcium binding sites in the former, segments 99–127 and 292–320, each 29 residues long. The similarity between human myeloperoxidase and bovine lactoperoxidase for these regions amounts to 48% and 72%, respectively. Thus it can be inferred that the best candidate for calcium binding is the second region.

As far as the activity of the compared peroxidases is concerned, it is not surprising that lactoperoxidase, myeloperoxidase and eosinophil peroxidase, which have quite similar functions, have the highest similarity. However a question still remains unanswered: why are the natural substrates of these three enzymes different?

We are grateful to J. Levin and G. Hazé for their help with the preparation of the manuscript.

Note. While this paper was being prepared for publication, the sequence of a cDNA encoding the entire bovine lactoperoxidase was reported (Dull et al., 1990). The N-terminus of the enzyme from Sigma was reported to be blocked and suggested to be Asp101 (reading frame numbering) from our results (Maillart, 1989). The same C-terminus was found. Only two differences occur between the two sequences: Dull et al. found Asp at position 9 and Asn at 349, instead of Asx and Arg, respectively.

REFERENCES

- Aune, T. M. & Thomas, E. L. (1977) *Eur. J. Biochem.* **80**, 209–214.
- Bardsley, W. G. (1985) in *The lactoperoxidase system* (Pruitt, K. M. & Tenovuo, J. O., eds) pp. 55–88, Marcel Dekker, New York.
- Bidlingmeyer, B. A., Cohen, S. A. & Tarvin, T. L. (1984) *J. Chromatogr.* **336**, 93–104.
- Booth, K. S., Kimura, S., Lee, H. C., Ikeda-Saito, M. & Caughey, W. S. (1989) *Biochem. Biophys. Res. Commun.* **160**, 897–902.
- Borstein, P. & Balian, G. (1977) *Methods Enzymol.* **47**, 132–145.
- Carlström, A. (1969) *Acta Chem. Scand.* **23**, 171–213.
- Derwahl, M., Seto, P. & Rapoport, B. (1989) *Nucleic Acids Res.* **17**, 8380.
- Dull, T. J., Uyeda, C., Strosberg, A. D., Nedwin, G. & Seilhammer, J. J. (1990) *DNA Cell Biol.* **9**, 499–509.
- Henschen, A. (1986) in *Advanced methods in protein microsequence analysis* (Wittmann-Liebold, B., Salnikow, J. & Erdmann, V. A., eds) pp. 244–255, Springer-Verlag, Berlin.
- Hoogendorn, H., Prissens, J. P., Scholtes, W. & Stoddard, L. A. (1977) *Caries Res.* **11**, 77–84.
- Isozaki, O., Kohn, L. D., Kozak, C. A. & Kimura, S. (1989) *Mol. Endocrinol.* **3**, 1681–1692.
- Johnson, K. R., Nauseef, W. M., Care, A., Wheelock, M. J., Shane, S., Hudson, S., Koeffler, H. P., Selsted, M., Miller, C. & Rovera, G. (1987) *Nucleic Acids Res.* **15**, 2013–2028.
- Johnson, K., Gemperlein, I., Hudson, S., Shane, S. & Rovera, G. (1989) *Nucleic Acids Res.* **17**, 7985–7986.
- Kimura, S. & Ikeda-Saito, M. (1988) *Proteins Struct. Funct. Genet.* **3**, 113–120.
- Kimura, S., Kotani, T., McBride, O. W., Umeki, K., Hirai, K., Nakayama, T. & Ohtaki, S. (1987) *Proc. Natl Acad. Sci. USA* **84**, 5555–5559.
- Kimura, S., Hong, Y. S., Kotani, T., Ohtaki, S. & Kikkawa, F. (1989) *Biochemistry* **28**, 4481–4489.
- Libert, F., Ruel, J., Ludgate, M., Swillens, S., Alexander, N., Vassart, G. & Dinsart, C. (1987) *EMBO J.* **6**, 4193–4196.
- Lu, H. S., Klein, M. L. & Lai, P. S. (1988) *J. Chromatogr.* **447**, 351–364.
- Magnusson, R. P., Gestautas, J., Taurog, A. & Rapoport, B. (1987) *J. Biol. Chem.* **262**, 13885–13888.
- Maillart, P. (1989) Thesis, University Paris VII.
- Morishita, K., Kubota, N., Asano, S., Kaziro, Y. & Nagata, S. (1987a) *J. Biol. Chem.* **262**, 3844–3851.
- Morishita, K., Tsuchiya, M., Asano, S., Kaziro, Y. & Nagata, S. (1987b) *J. Biol. Chem.* **262**, 15208–15213.
- Nichol, A. W., Angel, L. A., Moon, T. & Clezy, P. S. (1987) *Biochem. J.* **247**, 147–150.
- Poulos, T. L. & Kraut, J. (1980) *J. Biol. Chem.* **255**, 8199–8205.
- Reiter, B., Pickering, A. & Oram, J. D. (1964) in *Microbial inhibitors in food* (Molin, N., ed.) pp. 297–305, Almqvist & Wiksell, Stockholm.
- Rombauts, W. A., Schroeder, W. A. & Morrisson, M. (1967) *Biochemistry* **6**, 2965–2977.
- Sakamaki, K., Tomonaga, M., Tsukui, K. & Nagata, S. (1989) *J. Biol. Chem.* **264**, 16828–16836.
- Sievers, G. (1980) *Biochim. Biophys. Acta* **624**, 249–259.
- Sievers, G. (1981) *FEBS Lett.* **127**, 253–256.
- Tarr, G. E. (1982) in *Methods in protein sequence analysis* (Elzinga, M., ed.) pp. 223–232, Humana Press, Clifton, NJ.
- Ten, R. M., Pease, L. R., McKean, D. J., Bell, M. P. & Gleich, G. J. (1989) *J. Exp. Med.* **169**, 1757–1769.
- Tenovuo, J. O. (1985) in *The lactoperoxidase system* (Pruitt, K. M. & Tenovuo, J. O., eds) pp. 101–122, Marcel Dekker, New York.
- Theorell, H. & Akesson, A. (1943) *Ark. Kemi Mineral. Geol.*, **17B**, 1–6.
- Venturelli, D., Bittenbender, S. & Rovera, G. (1989a) *Nucleic Acids Res.* **17**, 7987–7988.
- Venturelli, D., Shirsat, N., Gemperlein, I., Bittenbender, S., Hudson, S. & Rovera, G. (1989b) *Nucleic Acids Res.* **17**, 5852.
- Yokote, Y., Arai, K. M. & Akahane, K. (1986) *Anal. Biochem.* **152**, 245–249.

SUPPLEMENT

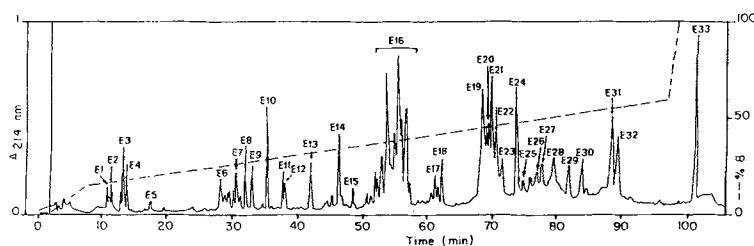


Fig. 1s. RP-HPLC of the endoproteinase Glu-C digest of Pe-A1. Column Nucleosil (SFCC, France; C_{18} , 5 μ m, 10 nm, 250 $\times$ 4.6 mm). Solvent A: 0.115% trifluoroacetic acid; solvent B: 80% CH_3CN , 0.1% trifluoroacetic acid. Column temperature: 40°C. Flow rate: 1 ml/min

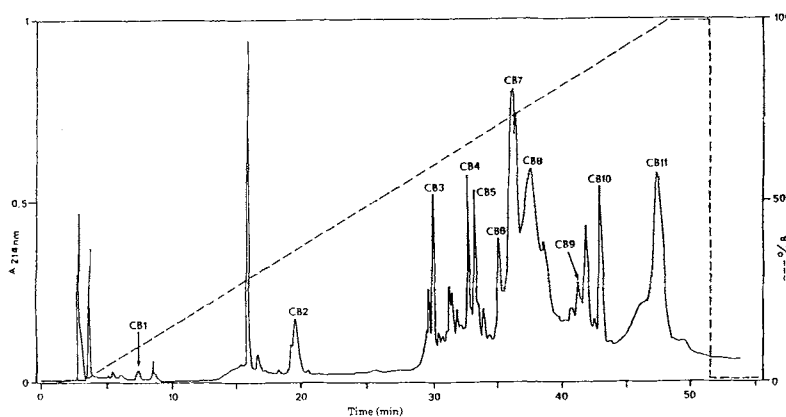


Fig. 3s. RP-HPLC of the CNBr digest of Pe-A1. Column: Nucleosil (SFCC, France; C_{18} , 5 μ m, 10 nm, 250 $\times$ 4.6 mm). Solvent A: 0.115% trifluoroacetic acid; solvent B: 60% CH_3CN /2-propanol (1/1, by vol.), 50 mM ammonium acetate buffer. Other conditions as in Fig. 1s

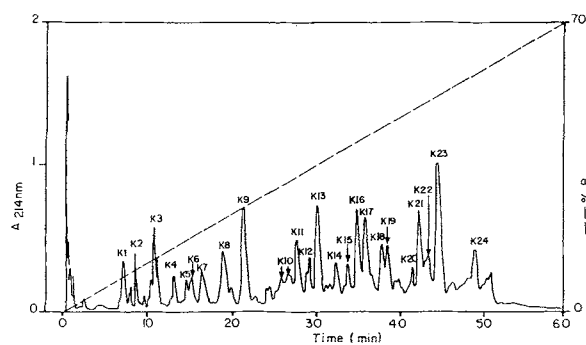


Fig. 2s. RP-HPLC of the endoproteinase Lys-C digest of Cm-A1. Column: Aquapore RP-300 (Brownlee; C_8 , 7 μ m, 30 nm, 30 $\times$ 4.6 mm). Solvent A: 0.115% trifluoroacetic acid; solvent B: 75% CH_3CN /2-propanol (1/1, by vol.), 0.1% trifluoroacetic acid. Other conditions as in Fig. 1s

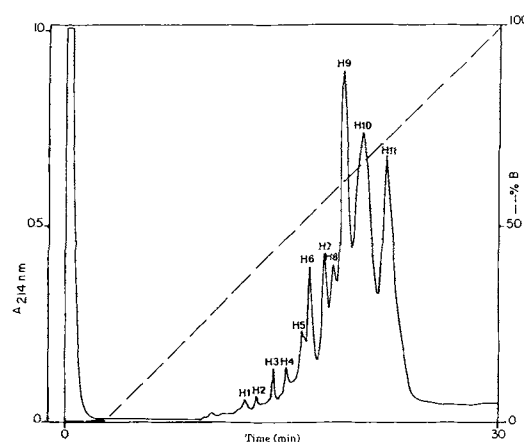


Fig. 4s. RP-HPLC of the hydroxylamine digest of Pe-A1. Column: Aquapore butyl (Brownlee; C_4 , 7 μ m, 30 nm, 30 $\times$ 4.6 mm). Solvent A: 0.115% trifluoroacetic acid; solvent B: 80% CH_3CN /2-propanol (1/1, by vol.), 0.1% trifluoroacetic acid. Other conditions as in Fig. 1s