UNIVERSITY of SOUTHAMPTON

Dept. of CLINICAL NEUROLOGICAL SCIENCES

THE PLASMA PROTEIN FETUIN

Common structural features of the mammalian fetuin family.

by William Michael Brown

A Thesis Submitted for the Degree of
Doctor of Philosophy, Faculty of Science,
University of Southampton.

March 1991

The moving finger writes, and, having writ,

Moves on: nor all your piety nor wit

Shall lure it back to cancel half a line,

Nor all your tears wash out a word of it.

From E. Fitzgerald's free
translation of
"The Rubáiyát of Omar Khayyám"

Contents.

****

Acknowledgments.

The list of people to whom I am indebted is long and I apologize in advance to anybody I have left out.

First and above all, I thank my supervisors: Prof. Norman Saunders, Dr. Richard Foreman and, more recently, Dr. Ian Day. Their support, scientific, moral and financial, is gratefully acknowledged.

My close colleagues Dr. Kate Dziegielewska and Dr. Mark Habgood also deserve considerable thanks. Continuing help, advice and the odd enzyme or microlitre of isotope from Dr. Keith Fox also deserves a mention.

Special thanks are due to Dr. B.A. Connolly for his permission to use the Applied Biosystems 381A and his and Dr. A. Worrall's patience in showing me how so to do, and to Dr. Mark Pickett for synthesising oligos at breakneck speed and very short notice, not to mention bringing them back to Connaught in the evening! Similarly special thanks are due to Drs. Jane Collins and David Garrod for patience, help and advice on the finer points of PCR and the use of the PCR machine, respectively.

Something should be said for my colleagues in The Connaught Hall Senior Common Room over the last three years (Mark Pickett, Helen Leigh, Anna Dickson, Brian Church, Derek Greer, Mike Carter, Andrew Shaw, Tahera Nensi and Sue Cooke) and the various JCR committee members and other staff at Connaught, but I have as yet not decided what! Special thanks are due to the current members of the SCR for their good grace (?) in accepting a few extra duties following my accident.

Thanks are also due to the numerous 3rd year project students, summer students, fellow postgraduates and assorted others who have contributed to the atmosphere in the lab.

On a more scientific note, I am indebted to all of our coauthors and collaborators, especially Dr. David Christie and Prof. Werner Müller-Esterl, whose visits to the lab. have been both scientifically profitable and greatly enjoyable.

The research reported in this thesis was funded by grants to Prof. Saunders, Dr. Dziegielewska and Dr. Foreman from the following bodies: Action Research for the Crippled Child, The Wellcome Trust, The Wessex Medical Trust, The Medical Research Council and The Sir Halley Stewart Trust.

Abbreviations.

20xSSC = 3M NaCl, 0•3M sodium citrate, pH 7•0.

20xSSPE = 3•6M NaCl, 2mM Na_2EDTA, 0•2M sodium phosphate, pH 7•7.

50xDenhardt's solution (Denhardt, 1966) = 1% (w/v) ficoll-400, 1% (w/v) bovine serum albumin, 1% (w/v) polyvinylpyrolidone.

2YT medium = 16g tryptone, 10g yeast extract, 10g NaCl per litre.

TAE = 40mM tris acetate, 2mM EDTA.

TBE = 90mM tris, 90mM boric acid, 2mM EDTA.

0

SM = 0•1M NaCl, 50mM tris-HCl, pH 7•5, 0•01% (w/v) gelatine, 8mM $MgSO_4 \cdot 7H_2O$.

UNIVERSITY of SOUTHAMPTON

ABSTRACT

FACULTY of SCIENCE

CLINICAL NEUROLOGICAL SCIENCES

Doctor of Philosophy

THE PLASMA PROTEIN FETUIN.

Common structural features of the mammalian fetuin family.

by **William Michael Brown**

The structure and tissue-specific expression of the bovine plasma glycoprotein fetuin and its homologues in other species has been examined. From the known partial amino acid sequence of bovine fetuin, degenerate oligonucleotide probes were designed. These probes were end-labelled using T4 polynucleotide kinase and were used to screen an adult bovine liver cDNA library by plaque hybridisation. A cDNA encoding the whole of the plasma protein was obtained and was fully sequenced on both strands. The deduced amino acid sequence was confirmed by protein sequence data in the literature [Christie et al., (1987), FEBS Letts. 214, 45-49]. The bovine fetuin cDNA was labelled by the random-primer technique and was used to screen a fetal sheep liver (pooled 40-60 day) cDNA library. A cDNA encoding the whole of sheep fetuin was obtained and was fully sequenced. The sheep fetuin cDNA was similarly labelled and was used to screen an adult pig liver cDNA library. An almost full-length pig fetuin clone was obtained and was sequenced. The deduced amino acid sequences were confirmed by amino-terminal protein sequence analysis by Dr. D.L. Christie of the glycoproteins purified from plasma. During this project it was reported that the protein encoded by rat clone pp63 [Auberger et al., (1989), Cell 58, 631-640] showed strong homology to bovine fetuin and human α_2-HS glycoprotein. Further analysis, reported in this thesis, demonstrates that clone pp63 encodes rat fetuin.

The three fetuin sequences reported in this thesis are compared with the other known members of the mammalian fetuin family: human α_2-HS glycoprotein, the rat protein encoded by clone pp63 and mouse fetuin. Sequence analysis reveals that fetuins comprise three domains: at the amino-terminus two cystatin-like domains and a unique carboxyl-terminal domain. A series of actual or potential sites for post-translational modification are apparent in the sequences.

By the techniques used (northern blot, RNAase protection assay and polymerase chain reaction) fetuin expression could only be detected in the liver. No fetuin mRNA could be detected in the fetal brain, although the protein can be immunocytochemically localised there. Data reported here show that in cattle, fetuin is a positive acute phase protein. It has been reported that human α_2-HS glycoprotein and rat fetuin are negative acute phase proteins.

1. Introduction.

This project was undertaken to examine the bovine plasma glycoprotein *fetuin* and fetuin-like proteins in other species. To this end, the cDNA and deduced amino acid sequences of bovine, ovine and porcine fetuin were determined and confirmed by protein sequence data.

It has been suggested that human α_2-HS glycoprotein is the homologue of bovine fetuin (Dziegielewska *et al.*, 1987; Christie *et al.*, 1987). This was confirmed by the complete sequence analysis and the fact that both proteins are encoded by single-copy genes in their respective genomes.

There has been a number of developments in the field during the course of the present project from other laboratories; it has been shown that the rat cDNA clone pp63 (Auberger *et al.*, 1989) displays strong homology to bovine fetuin and human α_2-HS glycoprotein (Haasemann *et al.*, 1991) and, further that it almost certainly encodes rat fetuin (Brown *et al.*, 1991a, see Discussion). The organisation of the rat fetuin (pp63) gene (Falquerho *et al.*, 1991) and the cDNA sequence of mouse fetuin (Yang *et al.*, 1991) have also recently been determined.

Analysis of the six fetuin sequences (human α_2-HS glycoprotein, the rat protein encoded by clone pp63, mouse fetuin and the three sequences presented in this thesis) indicates that mammalian fetuins are members of the cystatin superfamily as had previously been predicted on the basis of a partial sequence (Elzanowski *et al.*, 1988). The fetuins are characterised by a tripartite structure: two homologous sequences of cystatin-like structure ("cystatin domains") are positioned at the amino-terminus of the proteins and these are followed by a unique sequence segment ("terminal domain"), containing a proline-rich region. This three domain structure is exactly reflected in the organisation of the rat fetuin (pp63) gene (Falquerho *et al.*, 1991). The two cystatin domains, are each encoded by three exons and the terminal domain is encoded by a single exon.

Experiments were undertaken to examine whether or not *in situ* synthesis accounts for or contributes to the observed presence of fetuin or a fetuin-like protein in the developing mammalian forebrain. By the techniques used (northern blot, RNAase protection assay and PCR), it has not been possible to establish the presence of fetuin mRNA in the developing sheep brain.

Studies indicating that bovine fetuin is a positive acute phase protein are reported and discussed.

Much of the work reported in this thesis has been published or submitted for publication:

Dziegielewska, K.M., Brown, W.M., Casey, S.-J., Christie, D.L., Foreman, R.C., Hill, R.M., & Saunders, N.R. (1990) The complete amino acid and nucleotide sequence of bovine fetuin: its sequence homology with α_2-HS glycoprotein and its relationship to other members of the cystatin superfamily. *J. Biol. Chem.* **265**, 4354-4357

Brown, W.M., Christie, D.L., Dziegielewska, K.M., Saunders, N.R., & Yang, F. (1991) The rat protein encoded by clone pp63 is a fetuin/α_2-HS glycoprotein-like molecule, but is it the tyrosine kinase inhibitor PP^{63}? *Cell*, in press.

Dziegielewska, K.M., Brown, W.M., Gould, C.C., Matthews, N., Sedgwick, J.E.C., & Saunders, N.R. (1991) Variations in the serum concentration of fetuin-the bovine equivalent of human α_2-HS glycoprotein. *J. Comp. Physiol.*, submitted.

Brown, W.M., Christie, D.L., Dziegielewska, K.M., Nawratil, P., Saunders, N.R., & Müller-Esterl, W. (1991) The nucleotide and deduced amino acid structures of sheep and pig fetuin. Common structural features of the mammalian fetuin family. *Eur. J. Biochem.*, submitted.

Other work completed over the period of this project is reported in the following publications:

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., & Saunders, N.R. (1989) The nucleotide and deduced amino acid sequence of a γ subunit of bovine fibrinogen. *Nucleic Acids Res.* **17**, 6397

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., Saunders, N.R., and Wu, Y. (1989) The nucleotide and deduced amino acid sequence of sheep α1 antitrypsin. *Nucleic Acids Res.* **17**, 6398

Brown, W.M., Dziegielewska, K.M., Foreman, R.C. & Saunders, N.R. (1989) The nucleotide and deduced amino acid sequence of sheep serum albumin. *Nucleic Acids Res.* **17**, 10495

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., & Saunders, N.R. (1990) The nucleotide and deduced amino acid sequences of insulin-like growth factor II cDNAs from adult bovine and fetal sheep liver. *Nucleic Acids Res.* **18**, 4614

1.1. The plasma proteins.

The complement system has been reviewed extensively elsewhere (Müller-Eberhard, 1975; Reid & Porter, 1981; Campbell *et al.*, 1985; Law & Reid, 1988) and will not be considered further. Immunoglobulins (Putnam, 1987) and plasma lipoproteins (Scanu *et al.*, 1975; Scanu, 1987) similarly are reviewed elsewhere.

Historically, plasma proteins have been separated, and indeed classified and named, essentially on the basis of electrophoresis. The total concentration of plasma proteins in normal human blood lies in the range 5•7 to 8•0 g/dl (g/100ml) and the five classes, and approximate relative amounts, separated on the basis of free-boundary electrophoresis are albumin (54-58%), α_1-globulins (6-7%), α_2-globulins (8-9%), β_1-globulins (8-9%) and γ-globulins (11-12%). Each of these globulin fractions is now known to contain many individual proteins which could not be resolved by this technique. More modern techniques, including isoelectric focusing, immunoelectrophoresis and SDS polyacrylamide gel electrophoresis, have since revealed many plasma proteins, and, indeed, variants of them.

Plasma proteins as a whole will now be discussed with respect to 1) the coordinate control of gene expression, 2) the acute phase response and 3) the observation that some plasma proteins have been found in the developing mammalian brain. The origin of these brain proteins will be discussed along with a critical appraisal of the methods available and used to examine the observation. Finally, the plasma proteins *fetuin* and α_2-*HS glycoprotein* will be discussed individually with special reference to the structural and functional information available in the literature, to their presence and rôle in the developing brain, and to the relationship between these two proteins.

1.2. Coordinate gene expression.

Many studies of protein synthesis in general have used the plasma proteins as a system, simply because of ease of isolation, abundance, accessibility of material and considerable rate of synthesis. However, nothing has been found to distinguish the synthesis of plasma proteins by the liver from that of any other tissue and any other secretory protein. Plasma proteins are synthesised in the hepatocyte by ribosomes bound to the endoplasmic reticulum, (ER), cotranslationally pass into the lumen of the ER and enter the secretory pathway. The mammalian plasma protein cDNA sequences so far reported in the literature all predict fairly typical signal peptides to direct their secretion. The liver is the principal site of plasma protein synthesis (though as is discussed later, not the *only* site) and most of the research on plasma proteins has concentrated on hepatic expression. Hepatocytes are not restricted to the synthesis of any particular plasma protein; they can synthesise and secrete several simultaneously (Foucrier *et al.*, 1979; Courtoy *et al.*, 1981). Beyond this, several factors point to possible coordinate expression and this evidence will be described in the following paragraphs.

The chromosomal locations and the fact that many human plasma proteins are found in linkage groups throughout the genome has been extensively discussed and referenced by Bowman & Yang, (1987) and will only be mentioned briefly here.

It is clear from examination of Appendix I that the members of a number of protein families are tightly linked in the (human) genome. For example, the genes encoding serum albumin, α-fetoprotein and the group-specific component, Gc, which have evolved from a common ancestor (see Gorin *et al.*, 1981; Kioussis *et al.*, 1981; Alexander *et al.*, 1984; Dugaiczyk *et al.*, 1985; Yang *et al.*, 1985a,b) are present in a linkage group on human chromosome 4. The three fibrinogen genes, encoding the α, β and γ/γ' chains of the protein are closely linked on chromosome 4, see Appendix I for references.

Transferrin and a related protein, the melanoma antigen p97 (Brown *et al.*, 1982) are linked on human chromosome 3. Furthermore, two other metal-binding proteins (ceruloplasmin and α_2-HS glycoprotein) and the gene encoding the transferrin receptor are present in the same linkage group (see Appendix I and Bowman & Yang, 1987). Histidine-rich glycoprotein has also been assigned to chromosome 3 (van den Berg *et al.*, 1990).

Such functional and physical linkage of the genes suggests the possibility of coordinately regulated expression.

Recent studies on the control of plasma protein expression, made possible by the cloning and sequencing of plasma protein structural genes, have revealed highly conserved sequences in the proximal promoter regions. These conserved sequences are binding sites for *trans*-acting factors (*i.e.* sequence-specific DNA-binding proteins or transcription factors, for reviews see Dynan & Tjian, 1985; Maniatis *et al.*, 1987; Ptashne, 1988; Wasylyk, 1988; Johnson & McKnight, 1989, see also the section below on the acute phase plasma proteins and interleukin-6 (IL-6)).

These factors include those that are ubiquitous (*e.g.* nuclear factor 1 (NF-1), Borgmeyer *et al.*, 1984; Lichtsteiner *et al.*, 1987; Jones *et al.*, 1987; Raymondjean *et al.*, 1988 and Oct-1, Gerster *et al.*, 1987; Rosales *et al.*, 1987; Bohmann *et al.*, 1987; Fletcher *et al.*, 1987; Sturm & Herr, 1988), those that are liver-enriched or liver-specific (*e.g.* hepatocyte nuclear factor-1, (HNF-1) also known as LF-1B, see Schorpp *et al.*, 1987; Courtois *et al.*, 1987, 1988; Hardon *et al.*, 1988; Frain *et al.*, 1989; Baumhueter *et al.*, 1990, and CCAAT binding/enhancer-binding protein, (C/EBP), Landschulz *et al.*, 1988a, b, 1989) and a series of less common proteins, that by interacting with their respective elements, mediate the cellular response to phorbol ester treatment (Angel *et al.*, 1987; Lee *et al.*, 1987c), heat shock (Morgan *et al.*, 1987; Wu *et al.*, 1987; Sorger & Pelham, 1987, 1988), and serum stimulation (Gilman *et al.*, 1986; Prywes & Roeder, 1986; Treisman 1985, 1986, 1987). It is now clear that the many transcription factors are based on a limited

number of DNA-binding structures - the zinc finger, helix-turn-helix motif and the leucine zipper and basic region being the most common (see Latchman, 1990 for review).

By deletion and reporter gene analysis, it has been discovered that purified HNF-1 interacts with sequences in the promoter regions of the albumin, α-fetoprotein (AFP), α- & β-fibrinogen, α_1-antitrypsin and transthyretin genes (Cereghini *et al.*, 1987; Courtois *et al.*, 1987, 1988; Monaci *et al.*, 1988). Likewise C/EBP has been reported to bind to sites in the albumin, transthyretin and α_1-antitrypsin genes (Costa *et al.*, 1986; Cereghini *et al.*, 1987; De Simone *et al.*, 1987; Courtois *et al.*, 1987, 1988; Lichtsteiner *et al.*, 1987; Izban & Papaconstantinou, 1989; Maire *et al.*, 1989).

1.3. The acute phase response.

The acute phase response has been extensively discussed by Kushner (1982, 1988); Pepys & Baltz, (1983); Fey & Fuller, (1987); Schreiber, (1987); Schreiber *et al.*, (1989); Heinrich *et al.*, (1990); Fey & Gauldie, (1990) and will only be mentioned briefly here.

The *acute phase response* (Fig. 1) is the name given to the typical sequence of events in higher animals that follows a disturbance of homeostasis by, for example, inflammation, tissue injury, infection or neoplastic growth (for recent reviews see Schreiber, 1987; Heinrich *et al.*, 1990). The most notable effect of the acute phase response is the characteristic change in concentration of many plasma proteins, the increased level of fibrinogen being, quantitatively, most important in man. It is now clear that the observed changes in plasma concentration are largely as a result of changes in the rate of transcription from the respective genes. The early phase of the response is also marked by an increase in body temperature.

The acute phase plasma proteins - positive and negative - vary with species, and some apparently contradictory results have been obtained in transgenic mice experiments, making it very difficult to explain the rôle of these changes. For example transferrin (TF) is a negative acute phase protein in man, but a positive one in rodents. When the human TF gene was introduced into transgenic mice and an acute phase response induced, transcription from the endogenous mouse gene increased while that from the human transgene decreased (see Bowman *et al.,* 1990 for details and references, see also Discussion).

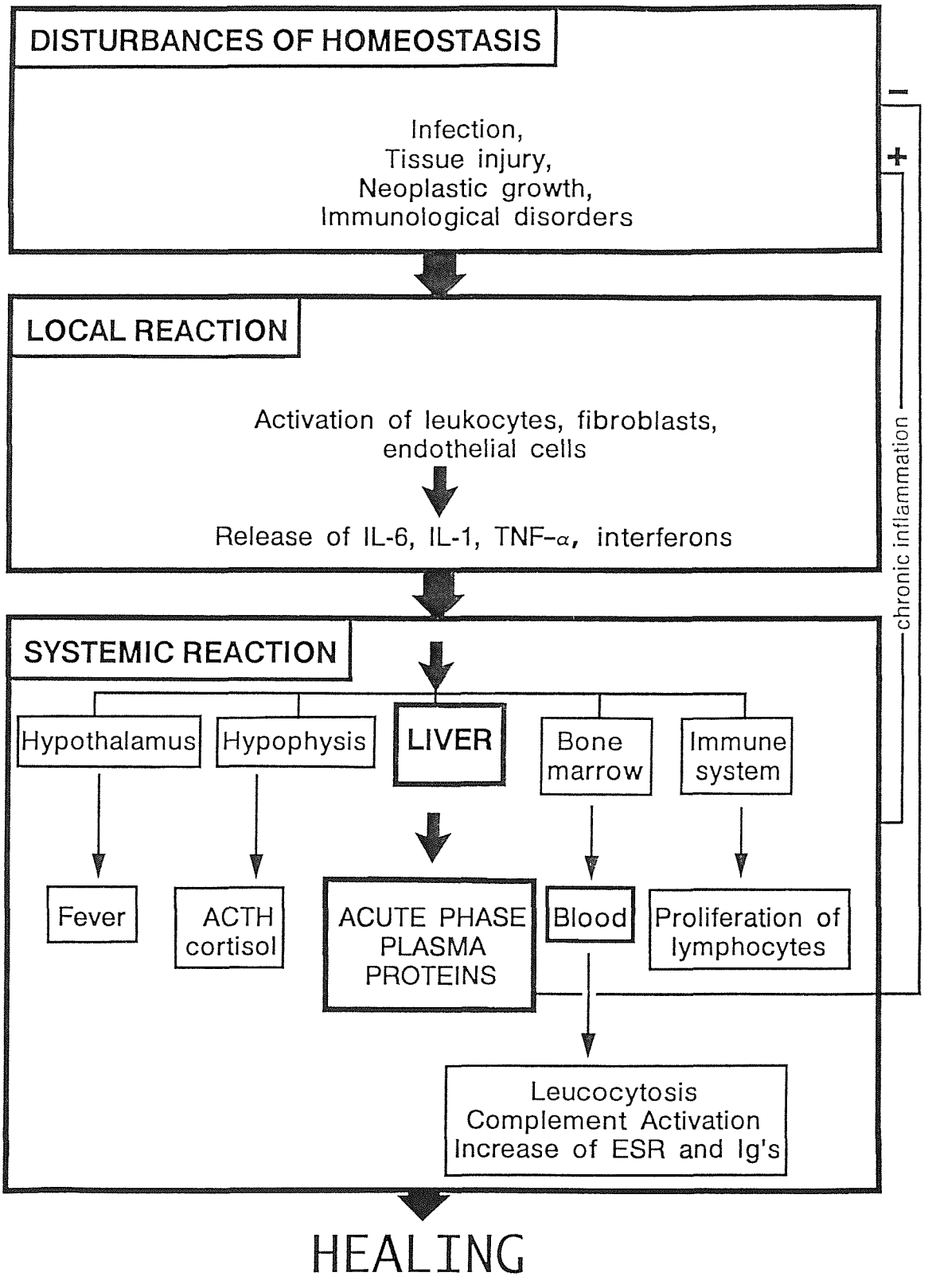
In most species plasma levels of fibrinogen, haptoglobin, α_1-antitrypsin and α_1-antichymotrypsin increase whereas the concentrations of albumin and transferrin decrease. Beyond this little is common; in man CRP (C-reactive protein) and serum amyloid A display dramatic increases, whereas in rat, the concentrations of T-kininogen (α_1-MAP, α_1-major acute phase protein, thiostatin), α_1-acid glycoprotein and α_2-macroglobulin show the largest rises (see Schreiber, 1987 for references).

The reaction comprises two components (see Fig. 1, redrawn from Heinrich *et al.,* 1990);
1) the *local reaction,* at the site of tissue injury which is characterised by blood clot formation, dilation of and leakage from blood vessels and, crucially, the local accumulation and activation of granulocytes and monocytes, which release acute phase cytokines (*e.g.* IL-6), resulting in,
2) the *systemic response,* marked by fever, leukocytosis, secretion of glucocorticoids and characteristic changes in the concentrations of the acute phase proteins.

Clearly the proteolytic breakdown of injured tissue implies the presence of active proteases. It has been noted that the concentration of many plasma antiproteases is increased in the response, and this is, perhaps, in order to limit damage to the rest of the body, and indeed the transport and communication systems based in the blood itself. Moreover, with regard to this damage limitation,

Fig. 1.

The Acute Phase Response



HEALING

Goldstein *et al.,* (1982) have suggested that the plasma concentration of ceruloplasmin is increased in the acute phase response to scavenge oxygen free radicals, $[O_2^{\bullet}]$, that activated leukocytes have been reported to produce (Goldstein *et al.,* 1975).

A further general heading for the function of the positive acute phase proteins is that of salvaging useful molecules released in the proteolytic breakdown. Clearly, the dividing line between mopping up molecules released in such breakdown and providing materials to regrowing tissue is a fine one and an increase in the concentration of a number of carrier plasma proteins may be to cater for this.

1.3.1. Acute phase response mediators.

Miller *et al.,* (1951) reported that the liver was the site of acute phase plasma protein synthesis and this fact leads to the inescapable conclusion that there must be messenger molecules released at the site of tissue injury to bring about the observed changes in plasma protein synthesis. As leukocytes gather at sites of tissue injury they have been intensively studied for any such messenger molecules that they might produce. A number of early experiments established that, indeed, leukocyte supernatants, when injected into laboratory animals, could bring about an artificial acute phase response (see Kampschmidt *et al.,* 1973; Powanda *et al.,* 1973; Wannemacher *et al.,* 1975). Beyond this, the work that led to the discovery and discrimination of the various factors produced by the leukocytes and their roles in the acute phase response is thoroughly reviewed and referenced by Heinrich *et al.,* (1990) and van Snick, (1990).

Thus, it has been established that monocytes and macrophages are the most important leukocytes with respect to the release of the factors involved. Human monocytes secrete interleukin-1 (IL-1) and tumour necrosis factor-α (TNF-α) and both were shown to have some effect on some acute phase proteins (see Heinrich *et al.,* 1990 for

references) but not to be the "*hepatocyte stimulating factor*" identified and characterised earlier by several groups (see Fuller *et al.*, 1987; Baumann *et al.*, 1987; Northoff *et al.*, 1987).

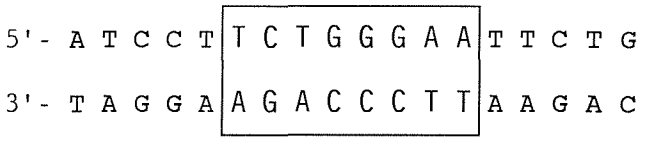
Further work concentrated on the interferons and led to the discovery that interferon β_2 was identical to the hepatocyte stimulating factor, and B-cell stimulatory factor-2, hybridoma plasmacytoma growth factor, and "the 26 kDa protein" (Billiau, 1986, 1987; Sehgal *et al.*, 1987). This molecule was rechristened interleukin-6 (IL-6) by Poupart *et al.*, (1987). See also Ann. N.Y. Acad. Sci. vol. 557 (1987) and van Snick, (1990).

The cloning of a number of acute phase protein gene sequences, and more particularly, the 5'-regulatory regions of such genes enabled the discovery of an IL-6 responsive element in the human haptoglobin (Oliviero *et al.*, 1987), rat α_1-acid glycoprotein (Prowse & Baumann, 1988) and rat α_2-macroglobulin (Gehring *et al.*, 1987; Tsuchiya *et al.*, 1987; Kunz *et al.*, 1989) genes.

Deletion mutants were constructed from the 5'-regulatory regions of such genes, ligated to appropriate reporter genes and transfected into suitable cell lines (*i.e.* liver-derived; Kunz *et al.*, (1989) used human hepatoma HepG2 cells). The expression of the reporter gene can then be monitored with respect to added IL-6.

By such deletion and reporter gene analysis, the element was narrowed down to a seven base sequence: 5'-CTGGGAA (Kunz *et al.*, 1989). Furthermore, that this sequence was genuinely responsible for the IL-6 response was tested by placing a synthetic consensus element *(vide infra)* upstream of a reporter gene (encoding chloramphenicol acetyl transferase, CAT) and assaying expression of the encoded reporter protein (Kunz *et al.*, 1989).

Fig. 2. The synthetic IL-6-responsive element that Kunz *et al.*, (1989) ligated onto a CAT reporter gene to confirm the consensus sequence first identified by Tsuchiya *et al.*, (1987).



Similar sequences have been located in the genes encoding rat α-, β- and γ-fibrinogens (Fowlkes *et al.*, 1984), rat T-kininogen (Fung & Schreiber, 1987), rat transthyretin (Fung *et al.*, 1988), human α_1-antitrypsin (Ciliberto *et al.*, 1985) and human CRP (Arcone *et al.*, 1988) See also Discussion.

More recent work has extended this discovery, and several types of IL-6 responsive elements have now been identified in acute phase protein genes (see Won & Baumann, 1990). The motif identified above is now known as a type B element (Won & Baumann, 1990). Fig. 3 shows the consensus sequences in the promoter region of a series of acute phase protein (positive and negative) genes.

More recently, two sequence-specific DNA binding proteins that have been shown to bind to these IL-6 responsive elements have been characterised and cloned (Akira *et al.*, 1990; Poli *et al.*, 1990).

Both are closely related to the transcription factor C/EBP, contain leucine-zipper structures and bind the same DNA consensus sequences. However, as pointed out by Poli *et al.*, (1990), while NF-IL6 expression was shown to be strongly induced by IL-6 (Akira *et al.*, 1990), activation of IL-6DBP occurred via a post-translational mechanism, suggesting that they are different members of a family of such proteins.

Fig. 3. Interleukin 6-responsive elements identified in acute phase plasma protein genes.

Abbreviations: HS, hypersensitive site; DRE, distal regulatory element; TTR, transthyretin; gp, glycoprotein; CRP, C-reactive protein, DE1, distal element 1.

The type A element has been shown to be IL-6-responsive in the gene encoding human hemopexin (Poli & Cortese, 1989). This sequence has been identified in many other acute phase genes (positive and negative).

The type B element has been shown to be IL-6-responsive, by deletion and reporter gene analysis, in the genes encoding the human haptoglobin (Oliviero *et al.*, 1987), rat α_1-acid glycoprotein (Prowse & Baumann, 1988) and rat α_2-macroglobulin (Kunz *et al.*, 1989) genes. It, also, has since been identified in many other acute phase genes (positive and negative).

Rat transthyretin (Fung *et al.*, 1988) is a negative acute phase protein, the other proteins shown here are positive acute phase proteins.

Gene	Location	Sequence Observed	Reference
Type A elements			
mouse TTR site 2	-1.8kb	A T T A G G A C A T	Costa *et al.*, (1986)
site 3	-1.8kb	G T T G A G T A A G	
albumin DE1	-105	A T T T T G T A A T	Lichtsteiner *et al.*, (1987)
albumin -3.5kb HS	-3.5kb	A T T G A G C A A T	Liu *et al.*, (1988)
rat α_1-acid gp DRE	21	A T T A A G A A A T	Won & Baumann, (1990)
	36	G T T G T G C A A T	
rat β-fibrinogen	-130	T T T G A G C A A C	Fowlkes *et al.*, (1984)
rat haptoglobin	-161	A T G A A G C A A G	Marinkovic & Baumann, (1990)
human CRP	-60	G T G G C G C A A A	Arcone *et al.*, (1988)
human hemopexin "A"	-117	G T G A T G T A A T	Poli & Cortese, (1989)
human haptoglobin "A"	-176	G T G A A G C A A G	Oliviero & Cortese, (1989)
"C"	-70	A T T A C G A A A T	
	CONSENSUS	T T N N G N A A T G	
Type B elements			
rat transthyretin	-323	A A C T G G G A A	Fung *et al.*, (1988)
rat α_1-acid gp	11	T G T T G G A A T	Won & Baumann, (1990)
	82	T T C T G G G A A	
rat β-fibrinogen	-152	T G C T G G G A A	Fowlkes *et al.*, (1984)
rat haptoglobin	-116	T A C T G G A A C	Marinkovic & Baumann, (1990)
rat α_2-macroglobulin	-166	T T C T G G G A A	Gehring *et al.*, (1987)
human CRP	-78	T G T T G G A A A	Arcone *et al.*, (1988)
human haptoglobin "B"	-127	T A C T G G A A A	Oliviero *et al.*, (1987)
	CONSENSUS	T T C T G G G A A T A T	

As C/EBP, IL-6DBP and NF-IL6 bind the same consensus sequence, it is possible that basal-level expression is controlled by C/EBP bound to the genes in question, and that in the acute phase response, C/EBP is displaced by one of these or other proteins - see also Discussion. However, how, or in conjunction with what, these proteins increase transcription from some genes, while decreasing it from others, remains unclear. The same consensus sequences are seen in the 5'-regulatory regions of positive and negative acute phase genes (Fig. 3). There is, obviously, more to acute phase gene expression than merely these conserved IL-6 responsive sequences. Furthermore, there are a number of reports in the literature of interactions between the factors mentioned here - IL-1, IL-6 and TNF-α - see Baumann *et al.*, (1987) and references therein.

Very little is known about the rôle of the negative acute phase proteins. Schreiber, (1987) suggested that perhaps the synthesis of albumin is decreased in the acute phase response as a "metabolic adaptation" to the increased requirement for the synthesis of other proteins; *i.e.* because of its quantitative importance in the liver, a decrease in its synthesis would release a significant amount of protein synthesising capacity. While this makes a degree of sense for albumin, simply in quantitative terms, the same could hardly be argued for the other negative acute phase proteins, bearing in mind their normal levels in plasma and the decreases observed. For example, Lebreton *et al.*, (1979) reported that the normal serum level of α_2-HS glycoprotein was 59•5∓0•02 mg/100ml (mean∓SEM, n=38), whereas the level in patients suffering from acute infectious diseases was 27•6∓0•02 mg/100ml (n=23).

1.4. Plasma proteins in the developing brain.

Following the seminal paper of Benno & Williams, (1978) many groups have reported the immunocytochemical localisation of plasma proteins in the developing mammalian brain. The presence in the developing brain of proteins, at least, immunologically related to

the plasma proteins, is now well established. At first sight the idea of plasma proteins being found in the brain is at odds with the concept of a blood-brain barrier, and a question that has, as yet, only been answered partially is the source of these plasma proteins. There is evidence for three possible origins in addition to that of *in situ* synthesis, which will be discussed in greater detail below. These three other origins are 1) uptake from the CSF, 2) "penetration" through the blood-brain and/or blood-CSF barriers and 3) uptake via nerve processes that extend beyond the blood-brain barrier, and these will now be briefly considered.

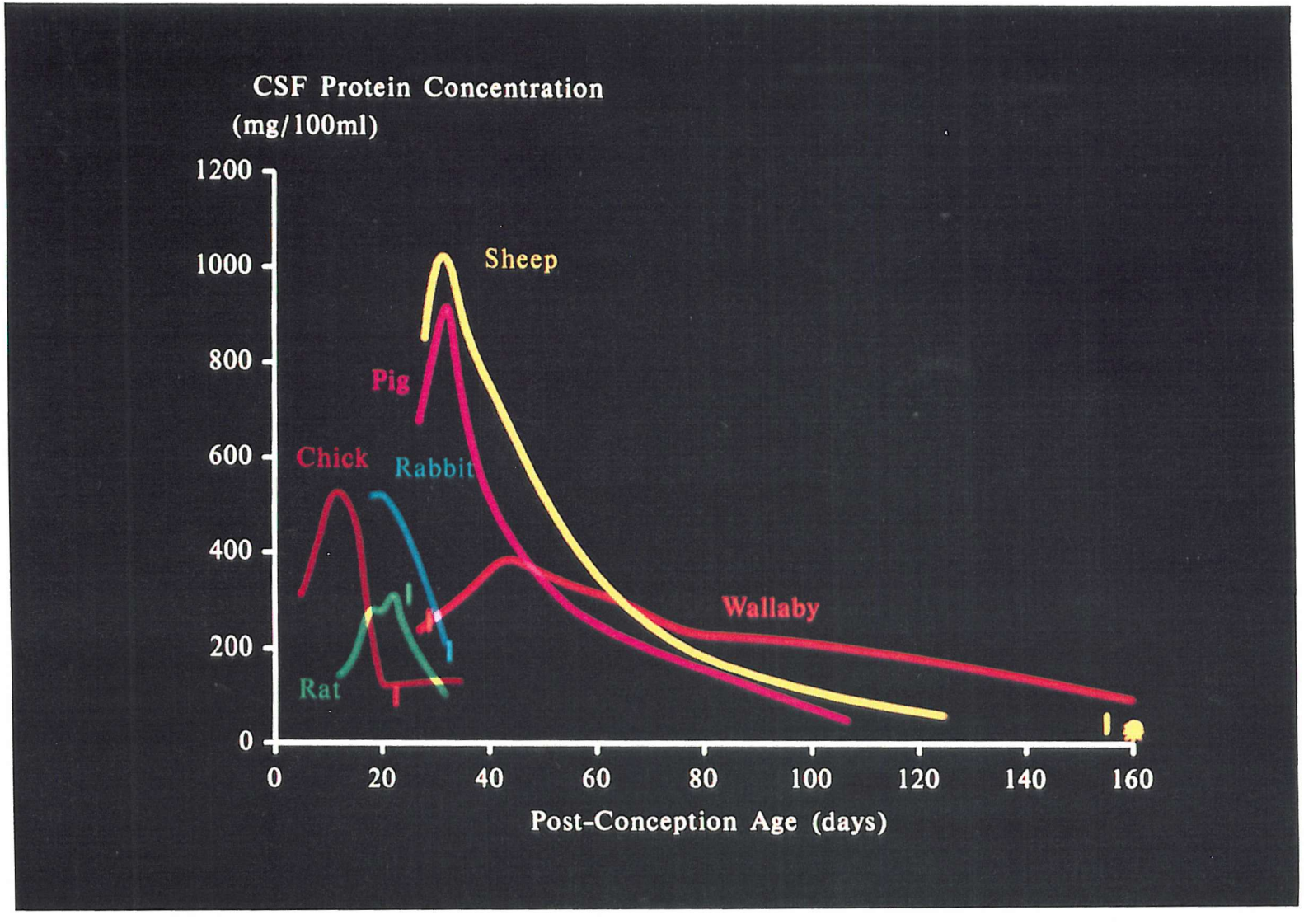
1.4.1. Uptake from the CSF.

Proteins immunologically indistinguishable from plasma proteins are present at high concentrations in fetal and neonatal CSF in many species, see Dziegielewska & Saunders, (1988) and Fig. 4. Note also that in Fig. 5, the precipitated CSF is strongly stained for fetuin. *In vivo* studies using chick embryos (Moro *et al.,* 1984) have shown that cells lining the mesencephalic cavity can take up foreign proteins (rat albumin and AFP). Cavanagh & Warren, (1985) reported that the presence of albumin within the cells of the developing rat forebrain can largely be attributed to uptake rather than *in situ* synthesis. *In vitro* tissue culture studies have shown uptake of AFP by brain hemisphere cells of mouse (Uriel *et al.,* 1981) and in the dorsal root ganglion cells of chicks (Hajeri-Germond *et al.,* 1984).

Fig. 4. Total protein concentrations in CSF through development.

Ordinate: total protein concentration (mg/100ml); abscissa: age from conception (days). The yellow asterisk indicates the mean adult value of total protein in CSF of all of the species indicated here, excluding chicken.

Figure redrawn from Dziegielewska & Saunders, (1988).


CSF Protein Concentration
(mg/100ml)
1200
1000
800
600
400
200
0
Sheep
Pig
Chick
Rabbit
Rat
Wallaby
0
20
40
60
80
100
120
140
160
Post-Conception Age (days)

1.4.2. "Penetration" across the blood-brain and blood-CSF barriers.

It has been reported that even at a very early age, the blood-brain and blood-CSF barriers are fully developed with respect to tight junctions and protein penetration and are not, contrary to popular belief, immature (Møllgård & Saunders, 1986). Indeed the blood-brain barrier can be seen to be intact in fetal brain sections immunocytochemically stained with antibodies to plasma proteins (see for example Reynolds *et al.*, 1987). Immunocytochemical staining is confined to the blood vessel lumen; there is apparently no "leakage".

1.4.3. Uptake via processes that extend beyond the blood-brain barrier.

Sparrow, (1981) reported the immunocytochemical localisation of plasma proteins in the neuronal perikarya and suggested that this was as a result of endocytosis and retrograde axonal transport in the same way that exogenous protein tracers have been shown to be taken up (for full discussion and references see Møllgård & Saunders, 1986). Broadwell & Brightman, (1976) reported that intravenously injected horseradish peroxidase was taken up by the endings of pre-ganglionic autonomic fibres, motor axons and sensory axons and was subsequently retrogradely transported to the neuronal perikarya. They suggested that it was via this route that blood-borne proteins and other high molecular weight substances might circumvent rather than breach the blood-brain barrier. Fink & Gainer, (1980) reported that an unidentified protein of 68 kDa is retrogradely transported in the axons of sciatic nerves in rat.

1.4.4. *In situ* synthesis.

There is evidence for the synthesis of a number of plasma proteins by the cells of the developing brain. It seems pertinent to point out at this stage that the idea of *in situ* synthesis is by no means

universally accepted or, indeed, in the light of evidence for other mechanisms, the universal explanation of the presence of plasma proteins in the developing brain. A frequently cited paper is that of Schachter & Toran-Allerand, (1982) which reported that albumin and AFP are not synthesised in the developing mouse brain. Both Darnell *et al.,* (1986) and Alberts *et al.,* (1983), citing the same work (Derman *et al.,* 1981, see also Powell *et al.,* 1984) state that α_1-antitrypsin, transferrin, transthyretin and albumin are synthesised in the liver and not in the adult brain or kidneys, on the basis of dot blot hybridisation experiments. However, it is crucially important to compare like with like; developmental stage, species and, perhaps most of all, the methods used must be kept firmly in mind when reading the literature. Derman *et al.,* (1981) state that transferrin is *not* synthesised in the adult mouse brain on the basis of dot blot hybridisations whereas Bloch *et al.* (1985) state that it *is* made in the adult rat brain on the basis of *in situ* hybridisation on tissue sections. There are rather too many variables to conclude that the papers contradict one another.

Early experiments in the field relied on incorporation of radiolabelled amino acids. Thus, Ali *et al.,* (1983) demonstrated that cells of newborn rat brain incorporate [^{3}H]-leucine into immunoprecipitatable α-fetoprotein and albumin in short-term culture.

A further general method used to demonstrate the presence of plasma protein mRNA in the brain is that of RNA extraction, translation *in vitro* or *in ovo* with a suitable radiolabelled amino acid, followed by immunological detection of radiolabelled protein. Using this approach, Dziegielewska *et al.,* (1985, 1986) reported that mRNA from developing rat cerebellum and human brain tissue respectively directed the *in vitro* synthesis of plasma proteins. Furthermore Møllgård *et al.,* (1988) reported that when human brain mRNA was translated in *Xenopus* oocytes, ceruloplasmin and transferrin could be immunologically detected *only* in the oocyte cytoplasm and not in the surrounding medium suggesting that the "brain" plasma proteins were not secreted. However, whether this is a genuine result or

some function of the kinetics of translation and secretion in oocytes remains to be rigorously proven - using for example a radiolabel pulse-chase experiment.

The fundamental weakness of these and other papers remains, however, the *immunological* detection of the product plasma protein(s), both from the point of view of the inability of a polyclonal antiserum to detect sequence differences (which might be the explanation of the above result) and a more general criticism of specificity and sensitivity. Sell *et al.,* (1985) made a devastating attack on the many papers that rely on immunoprecipitation of labelled proteins and set stringent criteria for such experiments.

More recently the problem has been tackled by hybridisation experiments - either dot or northern blot experiments using isolated mRNA or *in situ* hybridisation on tissue sections.

Clearly a necessary prelude to such experiments is the cloning and sequencing of the "liver" plasma proteins from the species and developmental stage of interest - thus within our group we have cloned and sequenced sheep α_1-antitrypsin (Brown *et al.,* 1989c), bovine γ' (γ_B) fibrinogen (Brown *et al.,* 1989a), sheep serum albumin (Brown *et al.,* 1989b), and bovine, sheep and pig fetuin (see Results). These clones and sequences are now available for the furtherance of this study, using, for example, *in situ* hybridisation. Many other plasma protein sequences from a variety of species have now been published.

The cDNA hybridisation approach was adopted by Levin *et al.,* (1984) and Dickson *et al.,* (1985). Cloned cDNAs were used as probes to perform northern blots and dot blots respectively. Dickson *et al.,* (1985) reported high levels of mRNA for transthyretin and transferrin in the developing rat choroid plexus, and a lower level of transferrin in the "rest of the brain". Levin *et al.,* (1984) reported that transferrin mRNA in rat brain increased from very low levels before birth to a plateau in the adult. Bloch *et al.,* (1985)

took this study further and performed *in situ* hybridisation using a cloned cDNA as a probe and reported that transferrin is synthesised in oligodendrocytes, "in most parts of the [adult rat] brain".

Recently Schreiber's group has published the results of a series of hybridisation experiments (Thomas *et al.,* 1989) using a number of plasma protein cDNA clones. Concentrations of mRNA were assessed using also dot blots and northern blots. Messenger RNAs encoding α_2-macroglobulin, transferrin, β_2-microglobulin, transthyretin, and ceruloplasmin were detected in brain RNA, by hybridisation to cloned (liver) cDNAs. However, the paper also reports that no mRNA was found in the developing rat brain at the ages studied for haptoglobin, α-fetoprotein, α_1-antitrypsin or T-kininogen, by the methods used.

Furthermore, Motojima & Goto, (1989, 1990) reported that in the rat choroid plexus, two transthyretin transcripts arise from dual promoters. In the liver, transcription is almost entirely from the proximal promoter, whereas in the choroid plexus, significant amounts of RNA are transcribed from the distal promoter (Motojima & Goto, 1990). In this work, perhaps, lies the molecular explanation of Schreiber's observation that expression of transthyretin by liver and choroid plexus appears to be differentially regulated. In particular, liver transthyretin expression has been reported to decrease markedly in the acute phase response *(vide infra),* whereas expression by the choroid plexus was apparently unaltered (see Dickson *et al.,* 1986).

However, as acknowledged by Motojima & Goto, (1990) this result is in direct conflict with that of Costa and colleagues who reported that in the choroid plexus and liver, transcription was from the same start point (see Costa *et al.,* 1989 and references therein).

From the various results in the literature it is clear that no other plasma protein mRNA is present in the brain (choroid plexus) in such abundance as the transthyretin message and that rather more sensitive and sophisticated techniques (*e.g. in situ* hybridisation

and PCR) will be required to continue this work, *i.e.* to examine whether any of the other plasma proteins found in the brain are synthesised there.

1.5. The plasma glycoprotein fetuin.

There has been considerable confusion in the literature with regard to the two fetal proteins fetuin and α-fetoprotein (AFP). The various protein, cDNA and genomic sequences now reported should really have put an end to this. While some reports in the literature clearly distinguished the two proteins as early as 1968 (Kithier *et al.,* 1968), as recently as 1989 (Nakai *et al.,* 1989), others confused the two and use the names interchangeably. Furthermore, a common undergraduate textbook (Smith *et al.,* 1983) states that "*Fetuin is the bovine plasma glycoprotein counterpart of* [human] *α-fetoglobulin* [AFP]".

For the record, fetuin and AFP are completely different plasma glycoproteins sharing no significant sequence homology. AFP shares a common origin with serum albumin and the group-specific component Gc (see Gorin *et al.,* 1981; Kioussis *et al.,* 1981; Alexander *et al.,* 1984; Dugaiczyk *et al.,* 1985; Yang *et al.,* 1985a,b) whereas fetuin is a member of the cystatin superfamily (see Discussion). AFP is present in fetal sheep and bovine plasma, as is fetuin. Fetuin is the bovine plasma glycoprotein counterpart of human α_2-HS glycoprotein, *(vide infra), not* of AFP.

Fetuin, a member of the α-globulin fraction of plasma proteins, is the predominant glycoprotein in fetal calf serum, reaching up to 50% of total plasma proteins in late gestation calves (Bergmann *et al.,* 1962; Dziegielewska, 1982). It was first purified and characterised from fetal calf serum by Pedersen, (1944). Since then, fetuin-like proteins have been identified in a number of mammalian species including sheep, pig, rat, mouse, goat, horse, cat and wallaby (Barboriak *et al.,* 1958; Marti *et al.,* 1973; Dziegielewska *et al.,* 1983; Jones *et al.,* 1988).

Immunocytochemical studies have demonstrated the presence of fetuin or fetuin-like proteins in the developing brain of many species. In particular fetuin or a related protein has been detected in the first cells of the cortical plate in the developing neocortex of such diverse species as sheep, pig, cow (Møllgård *et al.*, 1984, Reynolds *et al.*, 1987), marsupials (Jones *et al.*, 1988, 1991) and man (Dziegielewska *et al.*, 1987), see Fig. 5.

1.5.1. Structural & sequence data.

Beyond the initial report of Pedersen, (1944) little work was done on the protein until the early 1960's when Spiro and colleagues published a series of papers reporting protein chemistry and carbohydrate data. There was a number of reports in the 1940's and 1950's of fetuin or related proteins in a series of other species (see Dziegielewska, 1982 and Dziegielewska *et al.*, 1983 for references).

Fig. 5. Immunocytochemical staining for fetuin in the sheep brain.

A section from the lateral cerebral wall of a 35 day fetal sheep, was immunocytochemically stained using a rabbit-anti sheep fetuin antiserum (Dziegielewska *et al.*, 1980), and a modification of the indirect immunoperoxidase technique as described by Clausen & Thomsen, (1978). The section has been counterstained with haematoxylin. Thus, cell nuclei are blue, and the antibody reaction product is red-brown.

Note strong staining for fetuin in the cells of the cortical plate and choroid plexus, and in the CSF precipitated in the lateral ventricle. Similar results have now been obtained in many species (see text for references).

Figure taken from Møllgård *et al.*, (1984).

pia arachnoid

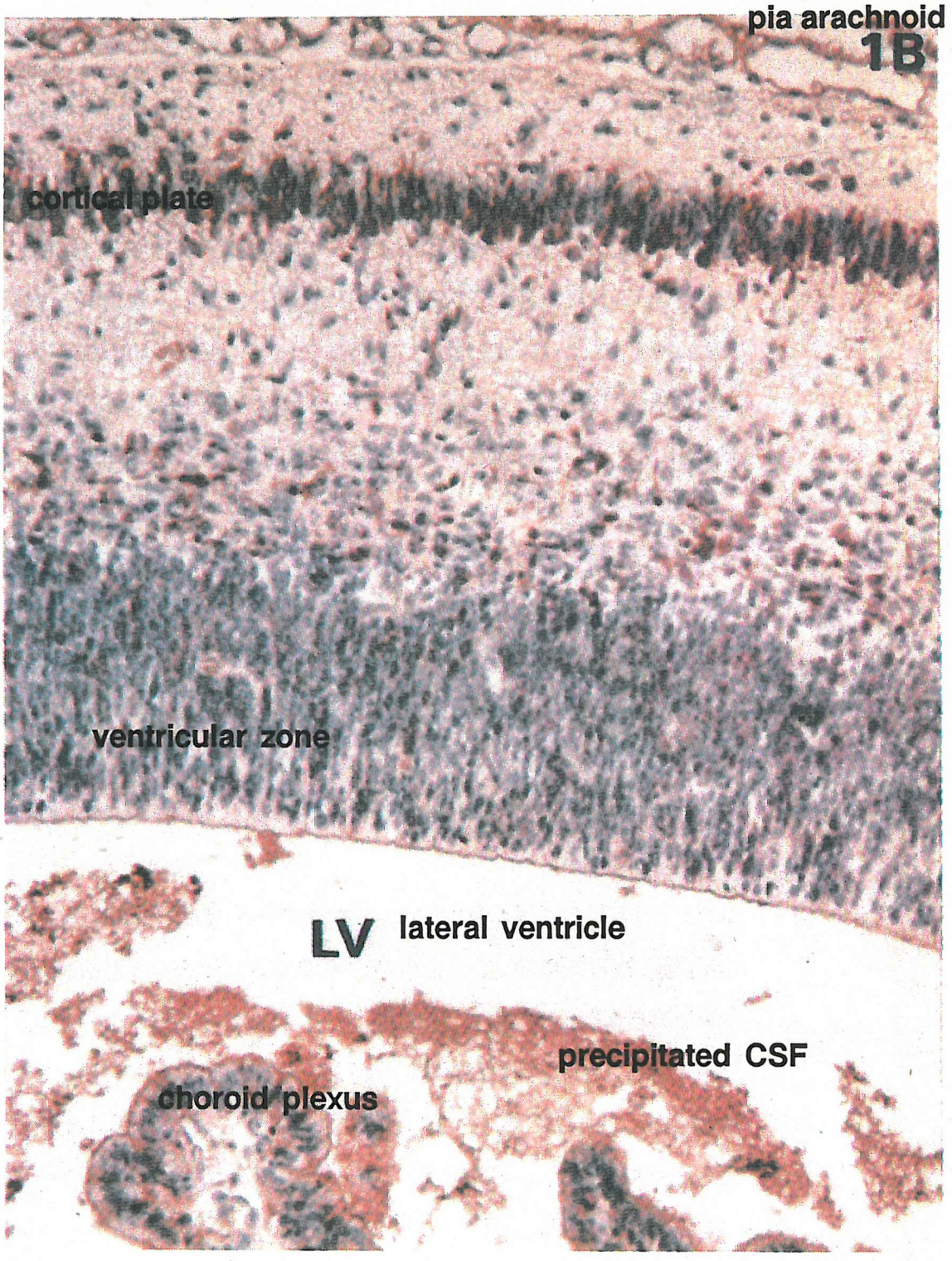
cortical plate

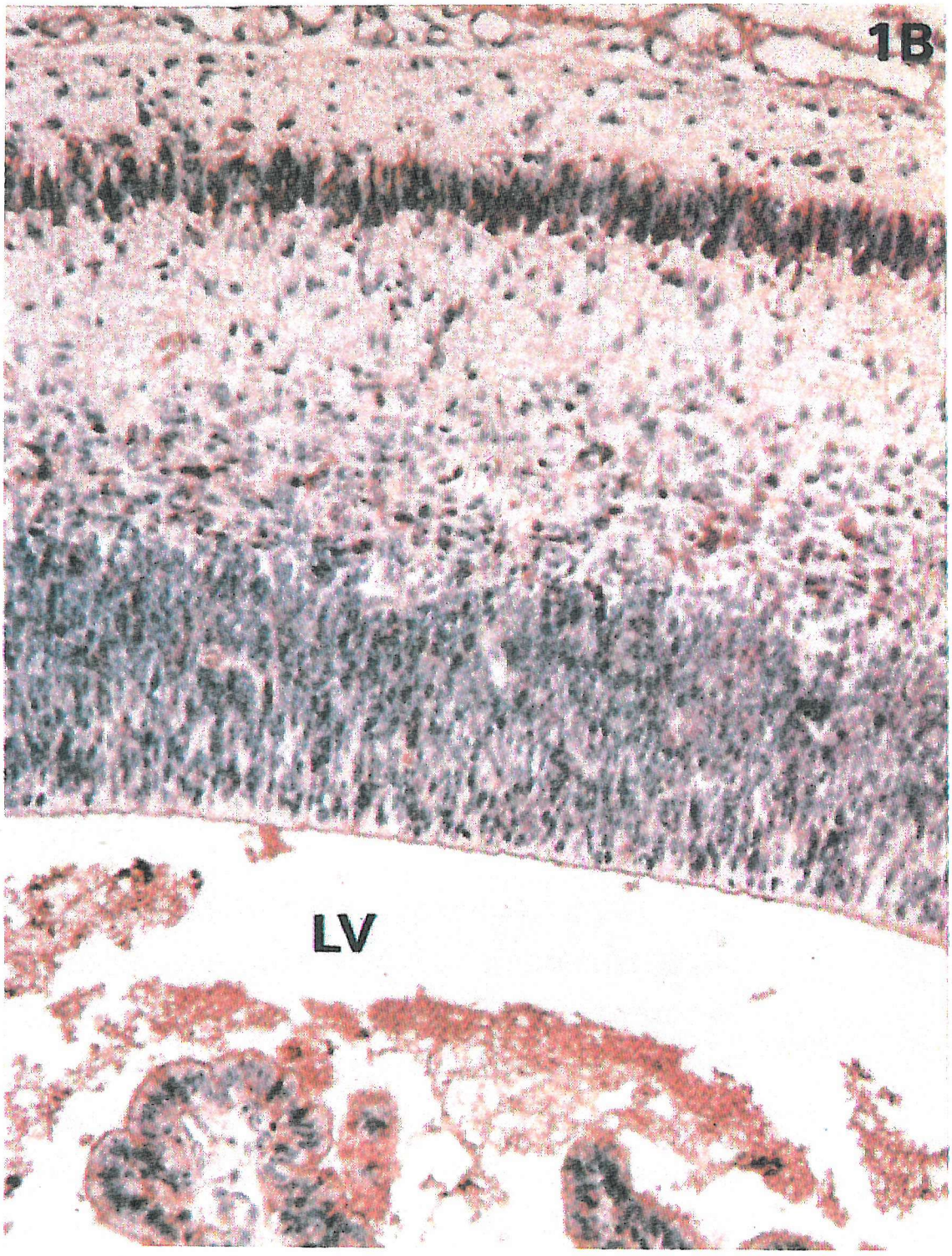
ventricular zone

lateral ventricle

precipitated CSF

choroid plexus


pia arachnoid
1B
cortical plate
ventricular zone
LV
lateral ventricle
precipitated CSF
choroid plexus


1B
LV

Spiro (1960) described a novel purification procedure, based on a low temperature ethanol fractionation, in the presence of Ba^{2+} and Zn^{2+} ions. This procedure yielded a fetuin of sufficient purity to conduct chemical and physical tests on the protein. Spiro, (1960) estimated the molecular weight to be 48400 Da, 26% of this being due to carbohydrates. From the earliest amino acid composition data (Fisher *et al.*, 1962; Spiro & Spiro, 1962) it was clear that bovine fetuin contained no methionine, twelve cysteine residues and a large excess of acidic groups, partly explaining the very low isoelectric point (pI) of 4 reported by Spiro, (1960).

Spiro, (1963) reported that the amino- and carboxyl-terminal residues were both isoleucine, that bovine fetuin was present in plasma as a single polypeptide chain, that although there were 12 cysteine residues, there were no free -SH groups, implying 6 intrachain disulphide bonds, and that a marked change in molecular shape occurred following reduction of the disulphide bonds.

Oshiro & Eylar, (1968, 1969) demonstrated that fetuin isolated by either the alcohol-metal ion fractionation procedure of Spiro, (1960) or the ammonium sulphate preparation of Fisher *et al.*, (1962) displayed microheterogeneity on analysis by starch gel electrophoresis at pH 4.2, and further demonstrated that this was largely due to terminal sialic acid residues on the carbohydrate structures.

A series of papers by C.M. Kay and colleagues, initially using optical rotation dispersion measurements (Green & Kay, 1963; Verpoorte *et al.*, 1963; Verpoorte & Kay, 1966) and later circular dichroism techniques (Murray *et al.*, 1969) examined structural properties of bovine fetuin. The results of these studies revealed that bovine fetuin had a low α-helical content ($\simeq$15%), and that while much of its conformation was determined by intrachain hydrogen bonding three other factors were involved. These were 1) interactions between hydrophobic regions of the protein, 2) the presence of six intrachain disulphide bonds and 3) the large overall negative charge as a result of terminal sialic acid

residues on the (six) carbohydrate moieties.

The primary site of fetuin biosynthesis is thought to be the liver (Thorbecke *et al.,* 1967), see also Results & Discussion. The amino-terminal sequence of bovine fetuin (purified by the method of Marti *et al.,* 1973) was reported by Alcaraz *et al.,* (1981). This was greatly extended by Christie *et al.,* (1987); their paper also contained internal sequence, determined by proteolytic dissection of the fetuin molecule.

1.5.2. Glycosylation.

Bovine fetuin has been used for many years as a prototypic glycoprotein, having 3 N-linked and 3 O-linked carbohydrate moieties. The structure of the carbohydrate groups of bovine fetuin has been extensively studied by many groups (see for example Begbie, 1974; Spiro & Bhoyroo, 1974; Baenziger & Fiete, 1979; Nilsson *et al.,* 1979; Oberholtzer *et al.,* 1981; Bergh *et al.,* 1983; Berman, 1986; Berman & Bendel, 1986; Townsend *et al.,* 1986; Takasaki & Kobata, 1986; Edge & Spiro, 1987; Berman *et al.,* 1988; Green *et al.,* 1988; Yet *et al.,* 1988; Cumming *et al.,* 1989; Bendiak *et al.,* 1989 and the many references therein).

1.5.3. Reported properties / functions of fetuin.

Numerous *in vitro* properties have been reported for fetuin preparations. However, when reading the literature it is important to consider the purification methods and the analyses used to demonstrate the purity of the protein; recently a number of the previously reported properties have been shown to be as a result of copurifying contaminant proteins (see for example Pierce *et al.,* 1979; Salomon *et al.,* 1982; Libby *et al.,* 1985; Zaitsu & Serrero, 1990).

Bovine fetuin is widely used as an additive in cell culture media (both in the form of fetal calf serum containing large amounts of fetuin and the purified protein itself) and the beneficial effects have been attributed to its growth-promoting properties (Puck *et al.,* 1968), although the detailed mechanism of this property remains unclear. Many other papers have reported effects of fetuin on cultured cells (see for example Fisher *et al.,* 1958; Ham, 1964; Marr *et al.,* 1962; Rizzino & Sato, 1978; Barnes & Sato, 1980; Gaillard *et al.,* 1985).

It has also been reported that fetuin purified by ammonium sulphate precipitation (Pedersen's method) contains appreciable amounts of α_2-macroglobulin (Marr *et al.,* 1962; Salomon *et al.,* 1982, 1984) and, bound both to the contaminating α_2-macroglobulin and to fetuin itself, platelet-derived growth factor (Libby *et al.,* 1985) which may explain the earlier results.

A number of papers have suggested that bovine fetuin is a serine protease inhibitor (see for example Galembeck & Cann, 1974; Rohrlich & Rifkin, 1981; Abdullah *et al.,* 1986). However, Dziegielewska, (1982) reported that sheep and pig fetuin did not display any trypsin-inhibitory activity.

Bovine fetuin was suggested to have lymphocyte-stimulating properties (Hsu & Floyd, 1976), to suppress T- and B-lymphocyte responses to phytohemagglutinin (PHA, Splitter & Everlith, 1982) to bind thyroid hormone (Fisher & Lam, 1974), to inhibit lymphocyte transformation (Yachnin, 1975), to bind lipids (Kumbla *et al.,* 1989), to stimulate lipogenesis (Cayatte *et al.,* 1990) and to stimulate differentiation of the adipogenic Ob17 cell line (Gaillard *et al.,* 1985). However Zaitsu & Serrero, (1990) recently described the purification of three adipogenic factors from Pedersen fetuin, again casting doubt on the earlier reported properties.

1.6. Human α_2-HS glycoprotein.

Human α_2-HS glycoprotein was discovered and purified from human plasma, independently, by two different groups using two different methods in 1960. Heremans, (1960) reported the isolation of a glycoprotein purified from the supernatant of plasma treated with zinc ions and called this protein α_2-Z-globulin.

Schmid & Bürgi, (1961) described a protocol for the purification of three low molecular weight α_2-glycoproteins from human plasma. In two preliminary steps, albumin and α_1-acid glycoprotein were removed by selective precipitation. Two of the three α_2-glycoproteins were then separated from the third by precipitation with Ba^{2+} ions. The co-precipitated proteins were called Ba-α_2-glycoproteins.

Later, Schultze *et al.*, (1962) realised the identity of the proteins separated by Heremans and Schmid & Bürgi, and proposed that the protein should be renamed α_2-HS-glycoprotein after its co-discoverers *H*eremans and *S*chmid. For reasons known only to J.P. Lebreton, and presumably those who reviewed the paper, it was renamed α_2-*SH* in Lebreton, (1977).

Human α_2-HS glycoprotein has been extensively studied ever since for four chief reasons:
1) it is highly polymorphic and has been used in forensic science;
2) its plasma level varies in a number of disease states;
3) it has been found at high levels in bone and dentine; and
4) there has been considerable dispute in the literature as to its structural properties, in particular whether it is present in plasma as a single polypeptide chain or two chains linked by a disulphide bridge. More recently, there has been further controversy as to the exact nature of the larger of the two chains. The work on human α_2-HS glycoprotein reported in the literature will now be discussed under these and other headings.

1.6.1. Purification.

Beyond the initial purification methods of Heremans, (1960) and Schmid & Bürgi, (1961), a series of protocols for the isolation of human α_2-HS glycoprotein has been reported (van Oss & Bronson, 1974; Lebreton, 1977; Gejyo & Schmid, 1981; Matsushima *et al.*, 1982; Arnaud *et al.*, 1983, 1988; Kellermann *et al.*, 1989). Furthermore affinity chromatography has been used, using antibodies to the protein (Lebreton *et al.*, 1979, Lewis & André, 1980).

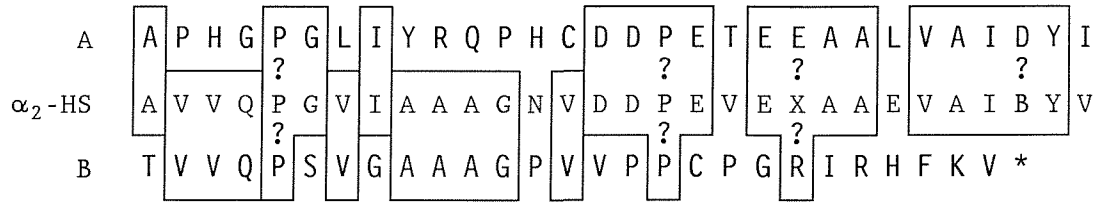
1.6.2. Structural and sequence data.

Even in the very earliest paper on the protein, it was suggested that human α_2-HS glycoprotein might comprise two polypeptide chains. Schmid & Bürgi, (1961) reported that two amino terminal *(ala, thr)* and two carboxyl terminal *(thr, leu)* amino acids were identified. However, at the time, the authors concluded that this was probably as a result of the Ba-α_2-glycoprotein fraction containing two different proteins. This was, indeed, found to be the case, but it was only part of the explanation. Further studies in Schmid's laboratory showed that the two Ba-α_2-glycoproteins could be separated, but that again two amino terminal *(ala, thr)* and two carboxyl terminal *(val, leu)* amino could be detected from *each* of the two proteins (Gejyo & Schmid, 1981). Amino acid analyses showed that the two proteins were very similar in having high cysteine and proline contents, but had slightly different arginine and histidine contents. From later results it seems possible that Schmid and colleagues had stumbled on the polymorphic nature of the protein *(vide infra),* as their proteins were purified from pooled human plasma. More recently, Schmid's group purified and sequenced the two chains of α_2-HS glycoprotein. The sequence of the (smaller) B chain revealed that *thr* and *val* were the amino- and carboxyl-terminal amino acids respectively (Gejyo *et al.*, 1983). It was later reported that *ala* and *leu* were the amino- and carboxyl-terminal amino acids respectively of the A chain (Yoshioka *et al.*, 1986), completing the explanation of the earlier results.

As mentioned above, however, there has been considerable disagreement as to whether α_2-HS glycoprotein exists as one or two chains. Lebreton *et al.,* (1979) stated most forcefully that "The protein consists of a unique polypeptide chain of about 50,000 daltons and has a unique amino-terminal residue, alanine." However, the authors go on to state that the protein was rather fragile and that "*spontaneous fragments*" were generated in some preparations. Lebreton *et al.,* (1979) also observed that these spontaneous fragments could be generated by serum proteases, and, *in vitro,* by trypsin. They further commented that a major polypeptide of M_r 30 kDa was generated. With the benefit of hindsight, it seems they had inadvertently discovered the trypsin sensitive site of α_2-HS glycoprotein (see Kellermann *et al.,* 1989 and Discussion).

Furthermore, Matsushima *et al.,* (1982) reported that human α_2-HS glycoprotein purified by their protocol from plasma was a single chain and that its migration was the same in SDS polyacrylamide gel electrophoresis under reducing and non-reducing conditions - which if true would argue against the idea of two chains linked by a disulphide bond. Matsushima *et al.,* (1982) also reported an amino-terminal sequence of the single chain. However, the amino-terminal protein sequence reported is only 50% correct in the light of later work. Moreover, analysis of the sequence they presented, bearing in mind the amino terminal sequences of the A and B chains since reported by Schmid's group, shows that it bears a disturbing similarity to a mix of the A and B chain sequences (Prof. W. Müller-Esterl, personal communication).

Fig. 6. Amino terminal sequence of human α_2-HS glycoprotein.
α_2-HS here = N-terminal sequence reported by Matsushima *et al.*, (1982), A & B = N-terminal sequences reported by Yoshioka *et al.*, (1986) & Gejyo *et al.*, (1983). Single letter code is used. X = unidentified amino acid, B = *asx*, that is *asn*, N or *asp*, D.



Thus the protein sequence reported by Matsushima *et al.*, (1982) could be a (rather poor) reflection on the state of the art of their protein sequencing, a result of some completely unconnected impurity, or as seems rather likely, in the light of the above sequence alignment, a reflection of the fact that they did, indeed, have both A and B chains of α_2-HS glycoprotein without realising it. A database search conducted with the sequence reported by Matsushima *et al.*, (1982) revealed no protein of greater homology than α_2-HS glycoprotein (data not shown).

As to the explanation of the observation that the migration of the purified protein was the same in SDS polyacrylamide gel electrophoresis under reducing and non-reducing conditions, there are a number of factors to be considered; 1) examination of Fig. 4, Matsushima *et al.*, (1982) shows that the migration was *similar* rather than *the same* under reducing and non-reducing conditions, 2) the (smaller) B chain reported by Gejyo *et al.*, (1983) has a molecular weight of only 5.4 kDa, would migrate virtually with the dye front, and might not stain particularly well, even if separated from it, and 3) it is not possible to predict the change in migration of any protein on reduction. It is now known that α_2-HS glycoprotein has a complex disulphide loop structure (Araki *et al.*, 1989; Kellermann *et al.*, 1989, see Discussion). A large change in migration could have arisen purely from the reduction,

and the presence or absence of the small B chain, especially bearing in mind the points raised above, might not have been noticed.

Even if it is now accepted that human α_2-HS glycoprotein is present in human plasma as two polypeptide chains held together by a single interchain disulphide bridge, more recently a further controversy has arisen, concerning the nature of the larger of the two chains. The principle site of α_2-HS glycoprotein synthesis is believed to be the liver (Triffitt *et al.,* 1976) and two groups reported the isolation of α_2-HS glycoprotein clones from liver cDNA libraries (Arnaud *et al.,* 1987; Lee *et al.,* 1987a). The complete nucleotide and deduced amino acid sequence of human α_2-HS glycoprotein has been determined (Lee *et al.,* 1987a). It revealed that the two chains (A & B) are encoded by a single mRNA transcript in the same reading frame, and that there is a 40-amino acid sequence between the A and B chain sequences apparently not seen in the mature protein. Unfortunately this 40 amino acid sequence became known as the "*connecting peptide*" (Lee *et al.,* 1987a). However, Kellermann *et al.,* (1989) reported that the "*heavy*" chain of α_2-HS glycoprotein (as opposed to the "*A*" chain of Yoshioka *et al.,* 1986) had at its carboxyl-terminus almost all of the connecting peptide, suggesting that the second cleavage is, in some way, an artefact of storage and/or purification, see Fig. 7.

A similar situation has been reported for tissue-type plasminogen activator (Pennica *et al.,* 1976) and the human plasma protein haptoglobin; in both cases the cDNA sequence encodes an arginine residue not seen in either chain of the mature protein which are linked by a disulphide bond (Yang *et al.,* 1983). The two chains of the mature proteins are again encoded in the same reading frame by a single mRNA transcript, there being a "connecting amino acid".

Fig. 7. Structure of human α_2-HS glycoprotein.

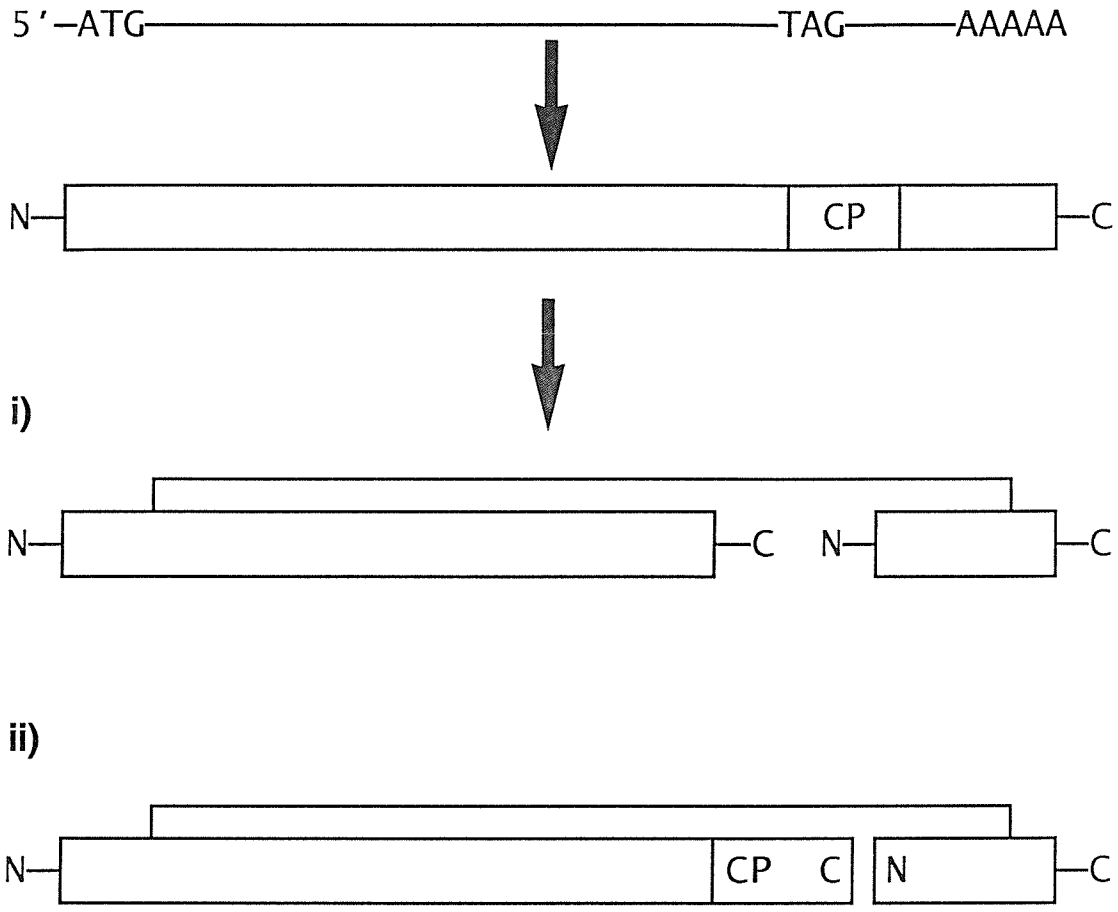


Fig. 7. The two chains of human α_2-HS glycoprotein are encoded by a single mRNA (Lee *et al.*, 1987a). Translation of this generates a precursor protein which is then proteolytically processed to remove the "*connecting peptide*" (CP) producing the A and B chains, the protein sequences of which were determined by Schmid's group (Gejyo *et al.*, 1983; Yoshioka *et al.*, 1986; i). More recently, however, it has been reported that α_2-HS glycoprotein in human plasma is present as "*heavy*" and "*light*" chains (Kellermann *et al.*, 1989; ii), the light chain being identical to the previously reported B chain and the heavy chain having at its carboxyl-terminus 39 of the 40 amino acids of the CP, suggesting that the second cleavage was an artefact of storage and/or purification.

The localisation of the α_2-HS glycoprotein gene (usually referred to as *AHSG* in genetic studies) to human chromosome 3 has been reported (Eiberg *et al.*, 1984; Cox & Francke, 1985; Arnaud *et al.*, 1987; Lee *et al.*, 1987a) and more recently, using *in situ* hybridisation, it has been assigned to 3q27→q29 (Magnuson *et al.*, 1988). It was observed that the α_2-HS glycoprotein gene on chromosome 3 was present in a linkage group with the ceruloplasmin and transferrin genes and a transferrin pseudogene - see Eiberg *et al.*, (1984); Bowman & Yang, (1987) and Appendix I. Histidine rich glycoprotein (HRG) has also been assigned to chromosome 3 (van den Berg *et al.*, 1990, see Appendix I).

1.6.3. The presence of α_2-HS glycoprotein in bone and dentine.

There is now a considerable literature on this subject. The organic matrix of bone contains a number of proteins, principally collagens. However, when bone is decalcified under conditions of neutral pH, the non-collagenous proteins can easily be separated from the largely insoluble collagen. Extracts prepared in this manner contain trace amounts of plasma proteins, simply from blood vessels within the tissue. However a number of plasma proteins are considerably concentrated within the mineral phase of bone, and are found at levels greatly in excess of that which could be explained by contamination. Furthermore, the concentration ratios of various proteins in bone are markedly different to those in plasma. In particular, albumin and α_2-HS glycoprotein have been reported by many authors to be present at high concentrations in bone matrix (see for example Ashton *et al.*, 1976; Triffitt *et al.*, 1976; Mbuyi *et al.*, 1982; Smith *et al.*, 1985 and references therein).

A series of early papers identified in rabbit and bovine bone homologous proteins, which became known as rabbit bone α-glycoprotein and G2B glycoprotein, respectively (Triffitt & Owen, 1973; Ashton *et al.*, 1974). Later work on human bone (removed during surgery) revealed a similar glycoprotein. The human protein was immunologically identified as α_2-HS glycoprotein (Dickson *et al.*,

1975; Ashton *et al.*, 1976; Triffitt *et al.*, 1976; Mbuyi *et al.*, 1982).

The binding of α_2-HS glycoprotein to calcium phosphates has been investigated by several groups. In each case, the basis of the experiment is to add varying amounts of Ca^{2+} (in the form of $CaCl_2$) and HPO_4^{2-} (in the form of Na_2HPO_4) to aliquots of serum. The mixture is then allowed to stand at 4°C overnight. The precipitate is then recovered, washed and EDTA-solubilised. Ca^{2+}, phosphate and the protein(s) of interest are then measured in 1) the starting serum, 2) the supernatant and 3) the EDTA-solubilised precipitate. Ashton *et al.*, (1976) and Smith *et al.*, (1985) both concluded that very low concentrations of calcium and phosphate could remove all α_2-HS glycoprotein from serum, demonstrating that, under these *in vitro* conditions, α_2-HS glycoprotein bound calcium and phosphate tightly. A number of other plasma proteins was examined - their concentrations were found to be unchanged by the procedure. Wilson *et al.*, (1977) reported that α_2-HS glycoprotein and albumin were present in fetal bone, at levels an order of magnitude greater than in adult bone, although the levels of the proteins in plasma changed little. A very similar result was reported by Quelch *et al.*, (1984) in neonatal bone. Furthermore Quelch *et al.*, (1984) reported a small number (n=5) of values, from the bone of children (aged 4-13 years), and these values, for albumin and α_2-HS glycoprotein, were intermediate, between the fetal/neonatal and adult levels. A series of pathological bone samples were also analysed; α_2-HS glycoprotein was found at a significantly higher level in the bone of patients with Paget's disease and *osteogenesis imperfecta* (OI), suggesting that there is no relationship between the plasma and bone levels *(vide infra).*

1.6.4. Changes in plasma levels.

As discussed above, the level of a number of plasma proteins alters in the acute phase response. Human α_2-HS glycoprotein is generally considered a negative acute phase protein (see Lebreton *et al.*,

1979 and Discussion). A series of cytokines are believed to be involved, but IL-6 is probably the major mediator of the acute phase response (for review see Heinrich *et al.*, 1990). Its involvement in the down regulation of human α_2-HS glycoprotein expression was confirmed by Daveau *et al.*, (1988). In the human hepatoma cell line HepG2, it was demonstrated that IL-6 and interleukin-1β caused a decrease in the expression of the α_2-HS glycoprotein, and that control was at the level of transcription (see Discussion).

The plasma level of α_2-HS glycoprotein has also been reported to be significantly reduced in Paget's disease of bone (Ashton *et al.*, 1976; Ashton & Smith, 1980), in children suffering from protein-energy malnutrition (Schelp *et al.*, 1980; Abiodun *et al.*, 1985; Abiodun & Olumu, 1987) and in patients with hypercalcemia (Crawford, 1984). However, in patients suffering from OI, plasma levels are significantly increased (Dickson *et al.*, 1983).

1.6.5. Reported properties/functions.

Besides the (suggested) involvement of α_2-HS glycoprotein in bone metabolism (see Colclasure *et al.*, 1988), and it being a negative acute phase protein *(vide supra)*, many other properties have been reported. These include possession of opsonic properties in bacterial phagocytosis by neutrophils (van Oss *et al.*, 1974), binding of DNA (Lewis & Andrē, 1978), promotion of endocytosis of radiolabelled DNA and latex particles by mouse macrophages (Lewis & Andrē, 1980), enhancement of macrophage phagocytic function (Lewis & Andrē, 1981), binding to lymphocytes transformed by Epstein-Barr virus (Lewis *et al.*, 1982; Lewis & Andrē, 1982), to be chemotactic for monocytes (Malone *et al.*, 1982) and to inhibit lymphocyte reactivity to phytohemagglutinin (Lewis & Andrē, 1984). See also Arnaud *et al.*, (1988).

1.6.6. Polymorphism.

The genetic polymorphism of human α_2-HS glycoprotein was first reported by Anderson & Anderson, (1977, 1979); on two-dimensional gel electrophoresis two allelic forms, first termed L and N were observed. Many techniques have now been used to separate the allelic forms (see Confavreux *et al.,* 1982; Cox & Andrews, 1983; Cox *et al.,* 1986; Umetsu *et al.,* 1984a,b; Weidinger *et al.,* 1984). More recently a "standard" method of thin-layer isoelectric focusing, followed by immunoblotting has been adopted by several groups (Boutin *et al.,* 1985; Yuasa *et al.,* 1985; Umetsu *et al.,* 1986; Sebetan, 1988; Sebetan & Heshmat, 1988; Thomas, 1989a; Fukuma *et al.,* 1990). Eighteen alleles have now been identified (Fukuma *et al.,* 1990) and from family studies, the mode of inheritance is autosomal codominant, *i.e.* the two alleles are independently expressed (for references see Appendix II). While *AHSG*1* has by far the highest frequency, because there are (at least) 17 other alleles, typically about 50% of a population have a phenotype other than the most common *AHSG 1-1.* Because of this the protein provides a useful genetic marker, which has been used extensively in forensic science (Westwood *et al.,* 1987b; Westwood, 1988; Thomas, 1989b).

1.7. The link between human α_2-HS glycoprotein and bovine fetuin.

From the partial fetuin sequence then available (Alcaraz *et al.,* 1981; Christie *et al.,* 1987) it was suggested by Christie *et al.,* (1987) that fetuin was the bovine homologue of the human plasma protein, α_2-HS glycoprotein, both chains of which had already been sequenced (Gejyo *et al.,* 1983; Yoshioka *et al.,* 1986). In the amino-terminal 105 amino acids, over 70% were in identical positions. The link had first been suggested by Dziegielewska *et al.,* (1987), who pointed out that the two proteins had a very similar amino acid composition (see Table I). This hypothesis was supported by the (circumstantial) evidence that in similar immunocytochemical experiments fetuin, α_2-HS glycoprotein or a related glycoprotein has been found in the developing cortex of all

species so far studied: human (Dziegielewska *et al.*, 1987), sheep (Møllgård *et al.*, 1984), cattle (Reynolds *et al.*, 1987) and rat (Sarantis & Saunders, 1986), see Fig. 5. A further fetuin-like protein and similar immunocytochemical staining has since been reported in the tammar wallaby (Jones *et al.*, 1988, 1991; Jones 1990). With the exception of the rat, the specific glycoprotein is expressed in the initial population of neurons that constitute the cortical plate. In the rat, α_2-HS glycoprotein appears later in brain development but in a site (Layer V-VI) in which the neurons originated from the first stage of cortical plate development (Brown & Reynolds, 1988).
However, despite the strong amino acid homology between fetuin and α_2-HS glycoprotein (Christie *et al.*, 1987) and the fact that it has been reported that both are members of the cystatin superfamily (Elzanowski *et al.*, 1988), several structural differences have been noted, particularly that antibodies to one protein do not easily cross-react with the other (Dziegielewska *et al.*, 1987; Christie *et al.*, 1987) and that α_2-HS glycoprotein has been reported to be present in plasma in a two chain form as discussed above (see Kellermann *et al.*, 1989) whereas bovine and ovine fetuin comprise single polypeptide chains (Spiro, 1960; Marti *et al.*, 1973).

Furthermore, to date, almost no common property has been demonstrated for both proteins, beyond their immunocytochemical localisation in the developing brain of their respective species. The only other property that both proteins have been reported to have is the ability to suppress lymphocyte responses to phytohem-agglutinin (Lewis & André, 1984, Splitter & Everlith, 1982).

Table I. Amino acid composition of fetuins[a]

Amino Acid	Calf fetuin[1]	Calf fetuin[2]	Lamb fetuin[3]	α_2-HS glycoprotein[4]	α_2-HS glycoprotein[5]	α_2-HS glycoprotein[6]
asx	33.4	33	32	31.6	30	32
thr	20.0	25	20	23.9	19	20
ser	25.1	26	22	22.4	18	20
glx	40.0	34	38	34.1	40	42
pro	34.2	34	35	43.7	39	41
gly	22.7	24	25	26.1	21	23
ala	37.4	33	40	33.8	36	38
val	40.9	40	38	41.5	33	35
met	0	0	0	2.9	1.2	1
ile	13.0	15	14	8.4	6	8
leu	27.1	27	28	33.8	28	30
tyr	8.4	7	8	4.8	6	7
phe	11.6	11	12	11.4	12	10
his	8.5	10	11	12.8	11	12
lys	14.1	16	16	15.4	16	16
arg	11.5	12	11	9.5	7	11
cys	11.6	12	12	11.0	7	11
trp	2	2	2	2.8	3	2

[a] in each case numbers given are residues per mol. protein.

[1] Alcaraz et al., (1981)
[2] Spiro & Spiro, (1960)
[3] Marti et al., (1973)
[4] Matsushima et al., (1982)
[5] Schmid & Bürgi, (1961)
[6] Heimburger et al., (1964)

As mentioned at the beginning of this Introduction, this project was undertaken to further examine bovine fetuin and fetuin-like proteins in other species, and to examine whether or not *in situ* synthesis accounts for or contributes to the observed presence of fetuin or a fetuin-like protein in the developing mammalian forebrain.

The cDNA and deduced amino acid sequences of bovine, ovine and porcine fetuin, other biochemical and structural experiments and studies indicating that bovine fetuin is a positive acute phase protein are reported and discussed. Experiments undertaken to establish whether or not fetuin is synthesised in the developing brain are also reported.

A major finding of this project is that the mammalian fetuins constitute a new family within the cystatin superfamily. The superfamily itself, and evidence for fetuin's membership of it are discussed in the Results and Discussion.

2. Materials & Methods.

2.1. Materials.

T4 DNA ligase was purchased from BRL. All other DNA-modifying enzymes were purchased from Boehringer or Pharmacia, unless otherwise indicated. Random primer labelling kits from Boehringer and Pharmacia were used. Radioisotopes and nylon membranes for library screening and northern blotting (Hybond N®) were purchased from Amersham International PLC. Dideoxynucleotide sequencing was performed principally using the Sequenase 2® reagents (U.S. Biochemicals Inc.), and [^{35}S]-α-dATP (Amersham, code SJ.1304, specific activity >1000Ci/mmol). The Taquence® kit (U.S. Biochemicals Inc.) and the Pharmacia T7 sequencing kit were also used. Charged nylon membranes, for Southern blotting, (Genescreen Plus®) were purchased from Dupont.

All other reagents were of "Analar" quality or better and were purchased from BDH, BRL, Fluka or Sigma. The pGEM®-3Zf(+) vector was purchased from Promega. SDS (sodium dodecyl sulphate) was from BRL (Ultra Pure). Glass-distilled phenol was obtained from Rathburn Chemicals Ltd., and was buffer-equilibrated according to Sambrook *et al.*, (1989). Organic solvents were Analar grade and purchased from BDH, except ethanol which was purchased from James Burrough (FAD) Ltd. and isoamyl alcohol which was from Sigma. Low melting point (LMP) agarose was purchased from BRL (Ultra Pure) or Sigma (molecular biology grade). Water was obtained from a Millipore Milli R/Q® water purifier. The source and grade of speciality reagents are indicated in the text.

2.2. Standard techniques.

Basic molecular biological and microbiological techniques were performed according to standard laboratory texts - Davis *et al.*, (1986); Miller, (1987); Ausubel *et al.*, (1989); Sambrook *et al.*, (1989). Antibiotics were stored and used according to Sambrook *et al.*, (1989). All microbiological techniques, genetic manipulation, radioisotope and chemical experiments were conducted with due

regard to the appropriate safety, containment and disposal rules.

2.2.1. Bacterial strains.

For genotypes and original references describing the bacterial strains used (*Escherichia coli,* strains Y1090, NM102, JM109, TG1, TG2), see Sambrook *et al.,* (1989).

2.2.2. Cloning vectors.

The cloning vectors used (λgt10, λgt11, pUC18/19, pGEM-3Zf(+), M13 mp18/19) are described in Sambrook *et al.,* (1989). The vector pSP64T and its modification (to create pSP64TW) is described later.

2.2.3. Lambda cDNA libraries.

The fetal sheep liver library in λgt10, was prepared by Dr. R.C. Foreman from pooled 40-60 day fetal sheep liver. The adult pig liver and adult bovine liver libraries, both in λgt11, were purchased from Clontech, California.

2.2.4. Blue-white screening - α-complementation.

In the pUC 18/19, M13 mp18/19, and pGEM3Zf(+) vectors the cDNA is ligated into a multiple cloning site (polylinker) or, in λgt11, the unique *Eco*RI site, located in the *lac Z* gene of the vector. This *lac Z* gene construct is derived from the bacterial *lac* operon and ordinarily encodes the amino-terminus of the (bacterial) β-galactosidase enzyme. Biosynthesis of this amino-terminal β-galactosidase fragment is induced (classically with the gratuitous inducer IPTG - isopropyl-β-thiogalactopyranoside) and the protein is capable of α- or intra-allelic-complementation with the defective β-galactosidase enzyme encoded by the bacterial chromosomal *lac Z* gene (in the

lac Z$^-$ strains used). Thus, when both fragments are produced, one from the plasmid / phage and one from the bacterial chromosome, a viable β-galactosidase enzyme results and it can convert the chromogenic substrate X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside, Horwitz *et al.,* 1964) into a blue compound facilitating identification of non-recombinant phage or plasmids. However, if a cDNA disrupts the plasmid or phage gene, a non-functional protein results which is unable to produce the blue compound from X-gal. Thus phage plaques or transformed bacterial colonies are white, if they contain an insert.

2.2.5. Labelling of probes.

The oligonucleotide probes used to obtain the initial bovine fetuin clone (Fig. 8 and Dziegielewska *et al.,* 1990) were labelled using T4 polynucleotide kinase (Boehringer) and [^{32}P]-γ-ATP (Amersham, code PB.10168, specific activity ≃3000Ci/mmol) and spun column-purified *(vide infra)*. Conditions used for the hybridisation and washing are described in Dziegielewska *et al.,* (1990). cDNA probes were labelled according to the random primer protocol of Feinberg & Vogelstein, (1983), as described in Brown *et al.,* (1991b), using [^{32}P]-α-dATP (Amersham, code PB.10204, specific activity ≃3000Ci/mmol) or [^{32}P]-α-dCTP (Amersham, code PB.10205, specific activity ≃3000Ci/mmol). Insert DNA was used either directly from a TAE LMP agarose gel or was phenol / chloroform extracted from a TBE LMP agarose gel, according to Sambrook *et al.,* (1989). Alternatively, insert DNA was purified from the gel using DEAE paper according to Sambrook *et al.,* (1989). The probes were purified from unincorporated nucleotides, using spun column chromatography.

2.2.6. Spun column chromatography.

Spun column chromatography was performed as detailed in Sambrook *et al.,* (1989) except that Sephadex G25-300® (coarse) resin in 2xSSC

was used and the glass wool was siliconized.

2.2.7. cDNA library screening.

Dilutions of the cDNA libraries were plated out in top agarose, using plating cells, according to the Amersham λgt10 and λgt11 cloning manuals. For initial screening (10^5-10^6 plaques) 22 x 22cm plates were used. Duplicate lifts were taken from each plate onto Hybond N nylon membranes (Amersham). These membranes were treated as per the manufacturer's instructions and were fixed by baking for 3 hours at 80°C without vacuum. Positive clones were picked, replated and rescreened on smaller plates (90mm diameter) at a much lower density, until single, at which point they were analysed by small-scale liquid lysate.

2.2.8. Prehybridisation of membranes.

Membranes were prewet in 2xSSC before being prehybridised at the hybridisation temperature for at least 3 hours. Hybond N membranes were typically prehybridised in 2xSSC, 0•1% SDS (w/v), 5xDenhardt's solution and 100μg/ml heat-denatured (10 minutes @ 100°C) herring sperm DNA (Sambrook *et al.*, 1989).

2.2.9. Hybridisation and washing.

Hybond N nylon membranes were hybridised overnight in a shaking waterbath typically in 2xSSC, 0•1% (w/v) SDS, 5xDenhardt's solution and 100μg/ml heat-denatured herring sperm DNA plus the labelled probe (≃2x10^6cpm/ml). Filters were washed twice at the hybridisation temperature for 20 minutes in 1ℓ 2xSSC / 0•1% (w/v) SDS, before being wrapped in Saranwrap®, and exposed in cassettes with intensifying screens (preflashed Fuji RX® film, -70°C).

Fig. 8. Design of the oligonucleotide probes used to obtain the bovine fetuin clone BFG.

N = A, C, G, T. X = A or G. Y = T or C. I = inosine (see Ohtsuka et al., 1985)
Sequence matches are indicated by an asterisk (*), inosine, which is neutral with regard to base pairing is indicated by a dash (-), mismatches are indicated by a gap.

64mer: 5'-d[TTXCAYTTIGTIGGXTCIACIACYTCYTTIGCIATXCAXTCIGTIGCIGCIACIGCXAAITCIAC]-3'

38mer: 5'-d[ATITCXCAXTCICCYTCIACIGCXTGYTGXTGYTGYTG]-3'

Protein sequence from which they were designed:

64mer, residues 181-202

```
protein sequence (single letter code)      v  e  f  a  v  a  a  t  d  c  i  a  k  e  v  v  d  p  t  k  c  n

DNA that could have encoded it          5'-GTN GAQ TTY GCN GTN GCN GCN ACN GAY TGY ATN GCN AAX GAX GTN GTN GAY CCN ACN AAX TGY AA-3'

antisense oligo made                    3'-CAI CTI AAX CGI CAI CGI CGI TGI CTX ACX TAI CGI TTY CTY CAI CAI CTX GGI TGI TTY ACX TT-5'
                                           **- **- *** **- **- **- **- **- *** *** **- **- *** *** **- **- *** **- **- *** *** **
DNA sequence found                      5'-GTG GAG TTT GCA GTG GCT GCT ACT GAC TGT ATT GCT AAA GAA GTC GTA GAT CCA ACC AAG TGC AA-3'
```

38mer, residues 86-98

```
protein sequence (single letter code)      q  q  h  q  h  a  v  e  g  d  c  d  i

DNA that could have encoded it          5'-CAX CAX CAY CAX CAY GCN GTN GAX GGN GAY TGY GAN AT-3'

antisense oligo made                    3'-GTY GTY GTX GTY GTX CGI CAI CTY CCI CTX ACX CTI TA-5'
                                           *** ***     *** *** **- **- *** **- *** *** **- **
DNA sequence found                      5'-CAG CAG ACG CAG CAC GCG GTG GAA GGA GAC TGC GAT AT-3'
```

2.2.10. Stripping of filters.

When necessary filters were stripped using the protocol recommended by Amersham - 0•4M NaOH @ 45°C for 30 minutes, followed by 0•1xSSC, 0•2M tris-HCl, 0•1% (w/v) SDS, pH 7•5, @ 45°C for 15 minutes. After this the filters were reprehybridised, rehybridised and rewashed as above.

2.2.11. Analysis of putative clones from library screening.

The method used was that of small-scale liquid lysate, by a mixture of protocols - Amersham λgt10 and λgt11 cloning manuals, Miller, (1987) and Sambrook *et al.,* (1989). The procedure was adopted in four situations.

1) as a test of the quality of a cDNA library - commercial or produced within the research group. A small proportion of the library was plated out, and 11 random clones and a no-phage control were processed to assess a) the percentage of clones containing inserts and b) the average size of these inserts.

2) as a method of preparing a high titre "stock" of the phage for long term storage (@ -20°C, -70°C and in liquid nitrogen). Following the lysis the cleared lysate was centrifuged and 4 x 1ml aliquots were placed in sterile cryovials. DMSO was added to 7% (v/v).

3) as part of the screening process - single positive clones were screened to decide which had the largest inserts and which were worth continuing to subclone and sequence.

4) in a scaled-up form as a preparative procedure. Selected positive clones for each protein were grown up to give a large amount of the purified insert for subcloning. Insert DNA was recovered by phenol / chloroform extraction from a preparative LMP agarose gel, according to Sambrook *et al.,* (1989).

In more detail, the protocol adopted was as follows. An overnight culture of the appropriate strain (NM102 for λgt10, Y1090 for λgt11) was prepared in 2YT medium to which maltose was added (final

conc. 0•4% (w/v)). A day culture was prepared from this in 2YT to which $CaCl_2$ was added (final conc. 5mM) and grown until $OD_{600} \simeq 0•25$. This culture was divided into aliquots - 6ml per sterile 50ml tube. A single λ plaque was cored into each tube. The tubes were shaken vigorously (250 rpm) at 37°C (NM102, λgt10) or 43°C (Y1090, λgt11) until the culture had completely lysed (5-6 hours). Chloroform (200μl) was added to each cleared lysate and the tubes were shaken for a further 20 minutes @ 37°C. The lysates were briefly centrifuged and the supernatant carefully removed, avoiding the chloroform and bacterial debris. A phage stock (1ml, stored @ -20°C) was taken at this point. A DNAase I / RNAase A digest was performed (enzymes @ 10μg/ml each, 37°C, 1 hour). After this, 5ml λ-PEG solution (20% (w/v) PEG-6000, 2•5M NaCl in SM) was added and mixed. The tubes were left on ice overnight. The agglutinated λ phage was pelleted (Beckman JA20.1 rotor, 12000 rpm, 4°C) and the supernatant discarded. Phage pellets were resuspended in 500μl SM, transferred to microfuge tubes, and vigorously shaken @ 4°C for 30 minutes. The tubes were briefly centrifuged and supernatants recovered. SDS & EDTA were added to final concentrations of 0•1% (w/v) and 5mM respectively and the tubes were heated to 70°C for 15 minutes to disrupt the phage protein coat. Once the tubes had cooled to room temperature, λ DNA was phenol/chloroform extracted and precipitated using isopropanol. Insert DNA was purified from the λ arms by preparative agarose gel electrophoresis, following *Eco*RI restriction.

2.2.12. Preparation of vector DNA for ligation of insert.

Vector DNA was digested to completion in the buffer appropriate to the restriction enzyme(s) in question before sodium acetate / ethanol precipitation. One of two further measures was also taken to reduce the background of non-recombinant transformants, though clearly in vectors where histochemical screening is available by α-complementation (pUC18/19, pGEM3Zf(+) and M13 mp18/19) this is of less importance.

When only a single restriction cut was made the digested DNA was

precipitated, treated with calf intestinal alkaline phosphatase (Northumbria Biologicals Ltd.), according to Sambrook *et al.*, (1989), phenol / chloroform extracted and reprecipitated.
Where the subcloning was to be forced (*i.e.* asymmetric), whenever possible, a triple digest was performed. In addition to the two enzymes required to generate the correct sticky ends to accept the insert, a "*wrecker*" enzyme was added - *i.e.* a polylinker restriction endonuclease with a site between the two sites of primary interest. If for example, the vector was digested with *Eco*RI and *Hind* III to accept an insert, then at the end of the digestion the excised polylinker itself constitutes an insert, has the correct sticky ends and, moreover, is present at a molecular ratio of 1:1 with the cut vector. It will, therefore, compete very effectively with the desired insert in the ligation mix. If, however, a triple digest is performed using *Eco*RI, *Hind* III and, for example, *Bam*HI then the polylinker is cut into 2 pieces which then have incompatible sticky ends and will not compete with the desired insert. Thus the background of non-recombinant transformed bacteria is greatly reduced. A small aliquot of the linearised vector DNA was analysed by agarose gel electrophoresis to assess yield and purity.

2.2.13. Preparation of transformation-competent cells.

Transformation-competent cells were prepared by the method of Mandel & Higa, (1970) as detailed in Sambrook *et al.*, (1989), except that the calcium chloride solution used was 100mM $CaCl_2$, 10mM tris-HCl, pH 8·0. The cells were stored overnight on ice in 200µl aliquots before use (Dagert & Ehrlich, 1979).

2.2.14. Transformation of competent cells.

Plasmid DNA (purified or a ligation mix) in a maximum volume of 30µl was added to a 200µl aliquot of transformation-competent cells, and incubated on ice for at least 30 minutes. Cells were heat-shocked for 2 minutes @ 42°C (circulating waterbath). At

this point M13-transformed cells were plated out according to Miller, (1987) on a lawn of TG2 cells, with X-gal and IPTG to facilitate identification of recombinant phage plaques.

Plasmid-transformed cells were allowed to recover at this point, *i.e.* to express the protein necessary to confer resistance to the selection antibiotic; 1ml prewarmed (37°C) 2YT medium was added to the heat-shocked cells and the mix was incubated @ 37°C for 30 minutes. Fractions of the mix were then spread on antibiotic-containing plates.

2.2.15. Preparation of M13 single stranded DNA sequencing template.

Template DNA was prepared as per Miller, (1987). TG2 cultures (10ml) were infected with a single M13 plaque and shaken for 4 hours at 240rpm in 100ml conical flasks (@ 37°C). The final single stranded DNA pellet was dissolved in an appropriate volume of TE (10mM tris, 1mM EDTA, pH 8.0) and a small aliquot was analysed by agarose gel electrophoresis to assess yield, purity and the presence or absence of an insert by comparison with single-stranded DNA prepared from a "blue" (non-recombinant) plaque.

2.2.16. Large scale, high purity preparation of plasmid DNA.

From a 500-1000ml culture, a large scale alkaline lysis (Birnboim & Doly, 1979) was performed according to Miller, (1987), with the exception that after the initial precipitation of plasmid DNA, an RNAase digestion was performed (DNAase-free RNAase A, 10μg/ml, overnight, 37°C) before phenol/chloroform extraction and reprecipitation. Plasmid DNA was finally purified using a pZ523® spun column (5'-3' Inc., Philadelphia), see Zervos *et al.*, (1988).

The quality and intactness of the plasmid DNA was checked 1) by restriction analysis and ethidium bromide-stained agarose gel electrophoresis and, on occasions, 2) by direct sequence analysis.

The concentration, and thus quantity obtained, was determined by U.V. spectrophotometry (Sambrook *et al.*, 1989).

2.2.17. Preparation of plasmid DNA for sequencing.

Plasmid DNA was prepared either on a large scale as above or on a miniprep scale (1-5ml) and purified using Tip-20® columns according to the manufacturer's protocol (Diagen Inc., Germany). Because of the design of the pUC plasmids and the M13 vectors (see Yanisch-Perron *et al.*, 1985) the same sequencing primers can be used. Also in pUC, it is possible to use a "reverse" primer, on the other strand, enabling sequencing of both ends of the insert regardless of orientation. Plasmid sequencing was performed the denaturation protocol given in the Sequenase booklet - 2-3μg purified plasmid DNA in water was denatured for 30 minutes @ 37°C, in 0•2M NaOH, 2mM EDTA, before sodium acetate / ethanol precipitation. The DNA pellet was redissolved in 7μl water and used directly in the annealing reaction.

2.2.18. Dideoxynucleotide sequencing.

Purified single stranded or denatured double stranded template DNA was sequenced by the dideoxy chain termination method of Sanger *et al.*, (1977) using a modified T7 DNA polymerase (Tabor & Richardson, 1989) and [^{35}S]-α-dATP (Amersham). The universal oligonucleotide sequencing primers used were the *"-40 primer"*, supplied in the Sequenase kit (5'-d[GTTTTCCCAGTCACGAC]-3') the *"-20 primer"*, from BRL (5'-d[GTAAAACGACGGCCAGT]-3') and the pUC *"reverse primer"* (5'-d[AACAGCTATGACCATG]-3', Boehringer). Numerous sequence-specific oligonucleotide primers were used throughout the project. The products of the Sequenase reaction were separated by electrophoresis, on a 55cm denaturing (8•3M urea) 6% polyacrylamide gel (Sequagel chemicals), run on a Macrophor® 2010 apparatus according to the manufacturer's instructions, except that the gel was removed from the glass plate (which was *not* silane-treated),

fixed (30 minutes in 2ℓ 10% (v/v) acetic acid) and dried under vacuum onto filter paper, before exposure in a cassette (Fuji RX film, room temperature).

2.2.19. Bovine fetuin gene copy experiment.

Bovine genomic DNA (Sigma product No. D4764) was digested with restriction enzymes *Bgl* II and *Eco*RI. The DNA fragments were resolved on an agarose (0•8%) TBE gel. DNA was capillary-transferred overnight to a nylon membrane (Hybond N) using 10xSSC as the transfer buffer. The probe used for Southern analysis (Southern, 1975) was the *Bgl* II-*Eco*RI fragment of the full length bovine fetuin clone, see Figs. 14 & 15, and comprised the entire 3' untranslated region of the clone. This *Bgl* II-*Eco*RI fragment was force-cloned into pUC19 and grown up. Purified insert was labelled using the random primer technique of Feinberg & Vogelstein, (1983). The prehybridisation, high-stringency hybridisation (65°C, 1xSSC, 5xDenhardt's, 100μg/ml heat-denatured herring sperm DNA, 0•1% (w/v) SDS) and washing (@ 65°C in 1xSSC, 0•1% (w/v) SDS) were performed as described above. The wet membrane was wrapped in Saranwrap and exposed in a cassette with intensifying screens (preflashed Fuji RX film, -70°C).

2.2.20. RNA preparation.

RNA was prepared by a modification of the guanidinium isothiocyanate / ultracentrifugation method of Chirgwin *et al.*, (1979). All glassware and the percussion mortar were rendered RNAase-free by baking (overnight @ 200°C), all plasticware used was either autoclaved or sterile-single use and all water from a Millipore Milli R/Q water purifier was DEPC-treated (according to Sambrook *et al.*, 1989) and autoclaved three times.
Source and grade of the reagents used: guanidinium isothiocyanate (BRL, Ultra Pure), caesium chloride (BRL, Ultra Pure), tris (Trizma Sigma, molecular biology grade), Na_2EDTA (BDH, Analar), Sarkosyl

(sodium lauroyl sarcosinate, Fluka Microselect), acids (acetic acid, hydrochloric acid, BDH, Analar), β-mercaptoethanol (Sigma, molecular biology grade), sodium acetate (Sigma, molecular biology grade), SDS (BRL, Ultra Pure), chloroform & n-butanol (BDH, Aristar), ethanol (James Burrough (FAD) Ltd, ultra pure).
The guanidinium solution was prepared as follows: 100g guanidinium isothiocyanate was added to 100ml DEPC-treated water, 10•6ml 1.0M tris (pH 7•6), 10•6ml 0•2M EDTA (pH 7•6), 21•2ml 20% (w/v) sarkosyl and 2•1ml β-mercaptoethanol. The volume was adjusted to 212ml with sterile water. The solution was then passed through a 0•22μm filter (Falcon) and stored at 4°C in a Nescofilm®-sealed, foil-covered sterile bottle. Sheep tissue (fetal or adult) was dissected as quickly as possible, placed in sterile plastic tubes and snap frozen in liquid nitrogen. All tissue was stored under liquid nitrogen until required. Tissue was weighed and quickly returned to liquid nitrogen before RNA was prepared. Frozen tissue was pulverised in a liquid nitrogen-cooled percussion mortar before being quickly scraped into a beaker containing the guanidinium solution (at room temperature). It was then quickly passed through a 20ml syringe with no needle to disperse any large lumps. The sticky solution was passed 6 times each through 19- and 23-gauge needles to shear the cellular DNA. Caesium chloride was added (1g per 2•5ml solution) and dissolved by gentle mixing. This solution was then layered over a caesium chloride cushion (5•7M CsCl, 0•1M EDTA, ρ=1•71∓5% g/ml). RNA was pelleted through this cushion by ultracentrifugation (SW40.1Ti rotor, 30000rpm, 20°C, 17 hours). The supernatant was carefully aspirated until the DNA at the interface had been removed. The tube was then inverted to drain. The RNA pellets were redissolved in "resuspension buffer" (10mM tris-HCl, 5mM EDTA, 1% (w/v) SDS, pH 7•4). The solution was chloroform / n-butanol (4:1, v/v) extracted. The organic phase was back-extracted with an equal volume of fresh resuspension buffer. Aqueous phases were combined and RNA was sodium acetate / ethanol precipitated. The quantity obtained was determined by U.V. spectrophotometry (Sambrook *et al.*, 1989).
Total RNA was also prepared by the method of Chomczynski & Sacchi, (1987), using the RNAgents® kit according to the manufacturer's

protocol (Promega).
Where required, poly(A)$^+$ RNA was prepared from total RNA so isolated using the Poly(A)Qwik® mRNA purification kit according to the manufacturer's instructions (Stratagene).

2.2.21. Northern blotting.

RNAs to be analysed by northern blotting were separated on a denaturing (2•2M formaldehyde) agarose gel. The protocol and MOPS/EDTA buffer formulation used were those described by Gerard & Miller, (1986). The gel was not ethidium bromide-stained, and the buffer was not continuously circulated. Halfway through the experiment the buffer was mixed. The BRL high MW RNA ladder was used as a marker. The components of this ladder are produced *in vitro* from pieces of λ DNA, subcloned into a transcription vector. The transcripts have poly(A) tails and can be detected by probing a blot with random primer-labelled λ-*Hind* III fragments.

The RNA was capillary-blotted overnight onto a Hybond N nylon membrane (Amersham), using 20xSSPE as the transfer buffer and fixed by baking @ 80°C for 3 hours without vacuum. Prehybridisation and hybridisation were performed in 5xSSPE, 50% (v/v) deionised formamide, 0•5% (w/v) SDS, 5xDenhardt's solution, 100μg/ml heat-denatured herring sperm DNA, as recommended by Amersham. Hybridisations were performed overnight in a shaking waterbath. High-stringency washing was as recommended by Amersham;
2 x [1ℓ 2xSSPE, 0•1% SDS, 15 minutes] @ the hybridisation temp.,
1 x [1ℓ 1xSSPE, 0•1% SDS, 30 minutes] @ the hybridisation temp.,
2 x [1ℓ 0•1xSSPE, 0•1% SDS, 15 minutes] @ room temperature.
The wet membrane was wrapped in Saranwrap and exposed in a cassette with intensifying screens (preflashed Fuji RX film, -70°C). RNA blots were not stripped.

2.3. Polymerase chain reaction (PCR).

The PCR experiment as conducted here comprised three parts;

1) reverse transcription of RNA to form a 1st strand cDNA, using a sequence-specific primer (antisense),
2) the PCR-proper - using the thermostable Amplitaq® enzyme (Perkin Elmer-Cetus), the same antisense primer and a sense primer, and
3) analysis of the PCR products by Southern blot - probed with the full length sheep fetuin cDNA, random primer-labelled as above.

Oligo design and synthesis is discussed below. Fig. 9 shows the position and orientation (*i.e.* sense or antisense) of the oligonucleotides used in this study on the sheep fetuin sequence. Oligo concentrations were assessed by U.V. spectrophotometry and the following figures: 1 $OD_{260} \simeq 20\mu g/ml$, MW of an n-mer $\simeq$ n x 340. The conditions used in the various experiments were based primarily on those reported by Syvänen *et al.*, (1988); Kanno *et al.*, (1989); Orlandi *et al.*, (1989); Kawasaki & Wang, (1989); Sambrook *et al.*, (1989); Becker-André & Hahlbrock, (1989) and those suggested by Perkin Elmer-Cetus.

2.3.1. PCR oligonucleotide design.

From the sequence comparison between bovine, ovine and porcine fetuins and human α_2-HS glycoprotein (see Fig. 24 and Brown *et al.*, 1991b) and antisera cross reactivity data (see Dziegielewska *et al.*, 1990 for references) it seems likely that the principle immunogenic determinants of fetuin (and α_2-HS glycoprotein) lie at or near the carboxyl-terminus of the proteins - in the terminal domain, D3 (see Figs. 24, 25 and Table VI). Thus, if the protein immunocytochemically localised in the developing brain is not *identical* to the liver protein, it is most likely to share antigens, and sequence, at or near, the carboxyl-terminus. Pairs of PCR oligonucleotide primers were, therefore, designed to amplify this region, bearing in mind the "rules" of PCR oligo design set out by

Lowe *et al.*, (1990).

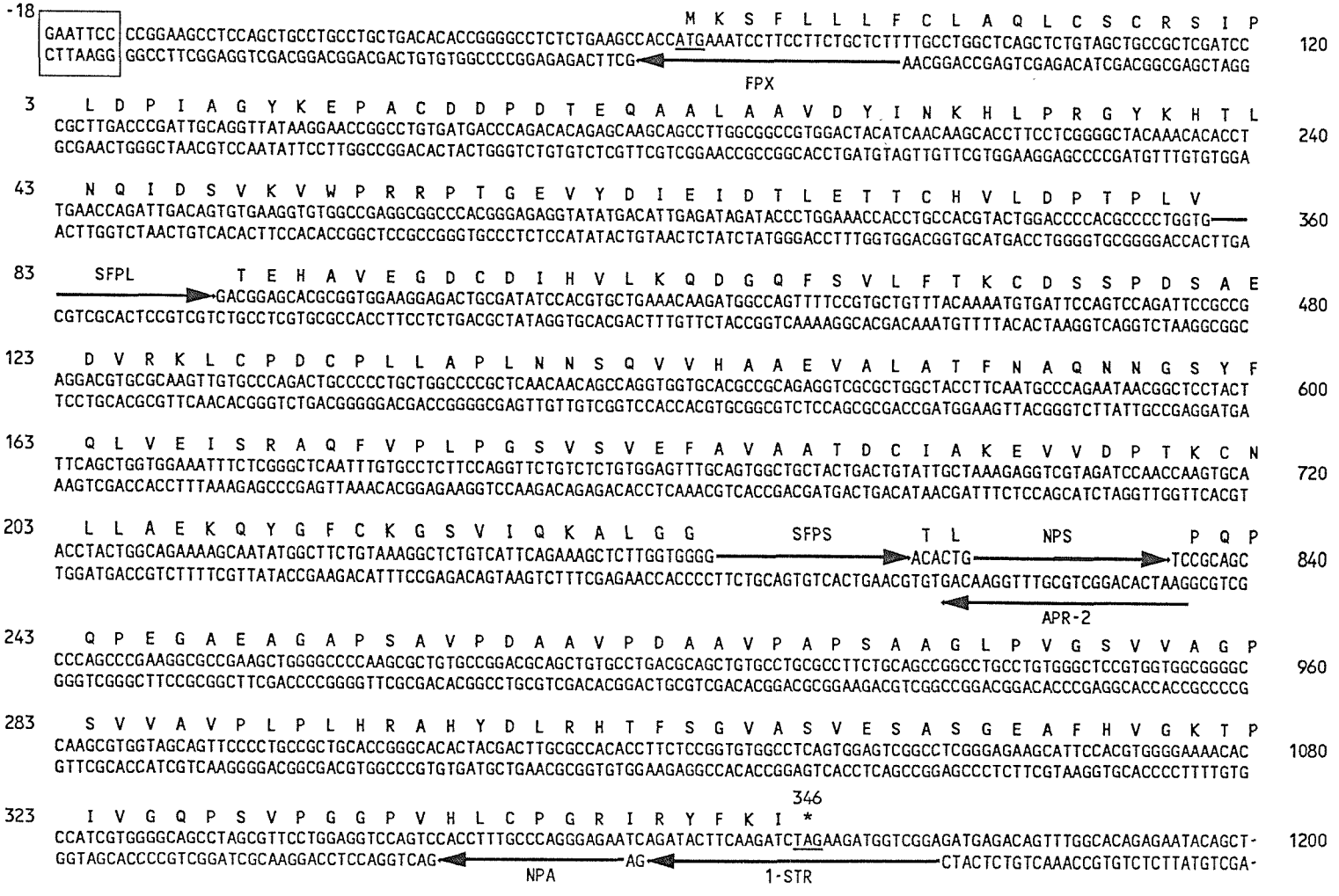
Name	Sequence
1-STR	5'-d[TCCGACCATCTTCTAGATCTTGAAGTATCT]-3'
SFPS	5'-d[AAGACGTCACAGTGACTTGC]-3'
SFPL	5'-d[AACTGCAGCGTGAGGCAGCA]-3'
NPS	5'-d[ggaattccTTCCAAACGCAGCCTGTGAT]-3'
NPA	5'-d[attcgaacaTTCTCCCTGGGCAAAGGTG]-3'
APR-2	5'-d[GGAATtcCAGGCTGCGTTTGGAACA]-3'
FPX	5'-d[AGAGCAGAAGGAAGGATTTCATGGTG]-3'

Table II. PCR oligonucleotide sequences. Perfect match sequence is in upper case letters. Lower case indicates extra or altered sequence to create restriction sites to facilitate force cloning of the product. Oligo FPX was made for a primer extension experiment and will also anneal to bovine fetuin mRNA. Oligos NPS and NPA are nested entirely within the 1-STR / SFPS oligo pair.

Oligo APR-2 was synthesised and HPLC-purified by Medprobe, Norway. Other oligonucleotides were synthesised using phorphoramidite chemistry (Caruthers, 1985) on an Applied Biosystems 381A (special thanks are due to Dr. B.A. Connolly, Dept. of Biochemistry, for his permission to use the 381A and his and Dr. A. Worrall's patience in showing me how so to do. Thanks are due to Dr. M.A. Pickett, Dept. of Microbiology, who also synthesised oligos used in this study).

Oligonucleotides were stripped from the column using "880" ammonia (BDH, Aristar) and precipitated. Concentrations were assessed by U.V. spectrophotometry.

Fig. 9. Position and orientation of the PCR oligonucleotides on the sheep fetuin sequence.



2.3.2. Reverse transcription.

Two sources of AMV-reverse transcriptase were used in the buffers recommended by the supplier;
1) Super RT® (Anglian Biotec), recommended buffer - 50mM tris-HCl, pH 8•3 (@ 20°C), 6mM $MgCl_2$, 40mM KCl.
2) Boehringer Mannheim, buffer 1 - 50mM tris-HCl, pH 8•5 (@ 20°C), 30mM KCl, 8mM $MgCl_2$, 1mM DTT.
Total RNA (10-20μg) or poly(A)$^+$ mRNA (2-5μg) was reverse transcribed in a reaction volume of 20 to 50μl containing the above buffers, 10 to 50pmol antisense oligo, 250μM dNTPs (dATP, dCTP, dGTP, dTTP diluted from 100mM stock solutions, Pharmacia), 30U RNAguard® (Pharmacia) or 40U RNAsin® (Promega) and 10 to 50U reverse transcriptase @ 42°C for 75 minutes. The reverse transcription mix was heated to 95°C for five minutes (Sambrook *et al.*, 1989) and quenched immediately in ice/water.

On a number of occasions, before the addition of the RNAase inhibitor or the reverse transcriptase, the mix was heated to 65°C for 3 minutes and cooled slowly (in an aluminium "hot block") to room temperature, to anneal the 1st strand oligo (Becker-André & Hahlbrock, 1989).

Where total RNA was used, the reaction mix was sodium acetate / ethanol precipitated and redissolved in sterile water before being split into the PCR mixes. Where poly(A)$^+$ mRNA was reverse transcribed it was divided directly into the PCR reaction mixes. PCR reaction mixes of 100μl volume were made up as per Perkin Elmer/Cetus' recommendation using the buffer supplied [10x stock has following composition: 100mM tris-HCl, pH 8•3 (@ 25°C), 500mM KCl, 15mM $MgCl_2$, 0•01% (w/v) gelatin] and Amplitaq (a recombinant thermostable Taq DNA polymerase). Sense and antisense oligos were added to a final concentration of 100 to 400nM. It was found empirically that 3•5mM $MgCl_2$ was optimal for the 1-STR / SFPS oligonucleotide primer pair. The reaction mix was overlaid with 100μl light mineral oil (Sigma) to prevent evaporation of sample, and kept on ice until loaded into the machine (15 to 30 minutes).

Magnesium chloride ($MgCl_2$) was purchased in the form of a sterile I•0M reference solution (Sigma, molecular biology grade) and was diluted appropriately with sterile water. The control DNA and oligonucleotide primers supplied with the Geneamp® kit were used when appropriate. However, as the starting point of the control experiment is carefully chosen, purified double stranded DNA and purified oligonucleotide primers, not a 1st strand cDNA / RNA mix, it was of very limited value as a control in this project. Indeed, the only things such an experiment truly "controls" are the nucleotide solutions, the buffer and the Amplitaq enzyme.

Special thanks are due to Dr. David Garrod, CRC Medical Oncology Unit, Southampton General Hospital, for his permission to use the Perkin Elmer DNA Thermal Cycler®, and to Dr. Jane Collins for her advice and patience.
Typically 35 rounds of amplification were performed; 60 seconds @ 94°C (denaturation), 60 seconds @ 65°C (annealing), 2 minutes @ 72°C (extension). The extension step of the final cycle was extended to 10 minutes @ 70°C and the samples were stored @ 4°C until collected. The PCR products were removed from the oil by extraction with 150µl chloroform (Sambrook *et al.*, 1989). Part (20%) of each PCR reaction was then analysed by agarose gel electrophoresis and Southern blotting as described below. The remainder was stored @ -20°C.

2.3.3. Southern blotting of PCR results.

This was performed using Genescreen Plus charged nylon membranes according to the manufacturer's instructions (Dupont), with the exception that the prehybridisation and hybridisation mixes contained 5xDenhardt's solution and heat-denatured herring sperm DNA (100µg/ml). Transfer was effected in 20xSSC, by capillary blotting, overnight. The transferred DNA was fixed to the membrane by simply allowing it dry at room temperature, as recommended by Dupont. The full-length sheep fetuin cDNA, SF6, was random primer-labelled and the probe was purified by spun column

chromatography. Prehybridisation and hybridisation were performed as detailed above @ 65°C in 2xSSC, 1% (w/v) SDS, 5xDenhardt's solution, 100μg/ml heat-denatured herring sperm DNA. Washing was carried out @ 65°C in 2xSSC, 0•1% SDS. The wet membrane was wrapped in Saranwrap and exposed in a cassette with intensifying screens (preflashed Fuji RX film, -70°C).

2.4. *In vitro* transcription and translation.

2.4.1. Modification of the pSP64T vector.

The plasmid vector pSP64T (Krieg & Melton, 1984) contains the 5'- and 3'-flanking regions of the *Xenopus laevis* β-globin mRNA, including a capping site, a ribosome binding site and a poly(A) tail. The purpose of the modification was to introduce an *Eco*RI site between the 5'- and 3'-flanking regions of the β-globin gene to facilitate the subcloning of the sheep fetuin and human α_2-HS glycoprotein cDNAs. Firstly, the existing *Eco*RI site was removed from the polylinker, using mung bean nuclease, according to Sambrook *et al.,* (1989). The vector was digested with *Bgl* II at a site between the 5'- & 3'-flanking regions of the globin mRNA and a *Bam*HI / *Eco*RI adaptor (Haymerle *et al.,* 1986, Fig. 11) was inserted.

Transformation-competent TG2 bacteria were transformed with purified pSP64T DNA (gift from Dr. R.C. Foreman) and transformed cells were selected on ampicillin plates (ampicillin @ 100μg/ml). A large scale plasmid preparation was performed as described above. Plasmid DNA (20μg) was *Eco*RI-digested to completion (as assessed by ethidium bromide-stained agarose gel electrophoresis) and the linearised DNA was sodium acetate / ethanol precipitated. The DNA was treated with mung bean nuclease (BRL) according to Sambrook *et al.,* (1989) and sodium acetate / ethanol precipitated. The ends of the DNA were polished using the Klenow fragment of DNA polymerase in H buffer (50mM tris-HCl, 10mM $MgCl_2$, 100mM NaCl, 1mM dithio-erythritol, pH 7•5) and the DNA was sodium acetate / ethanol

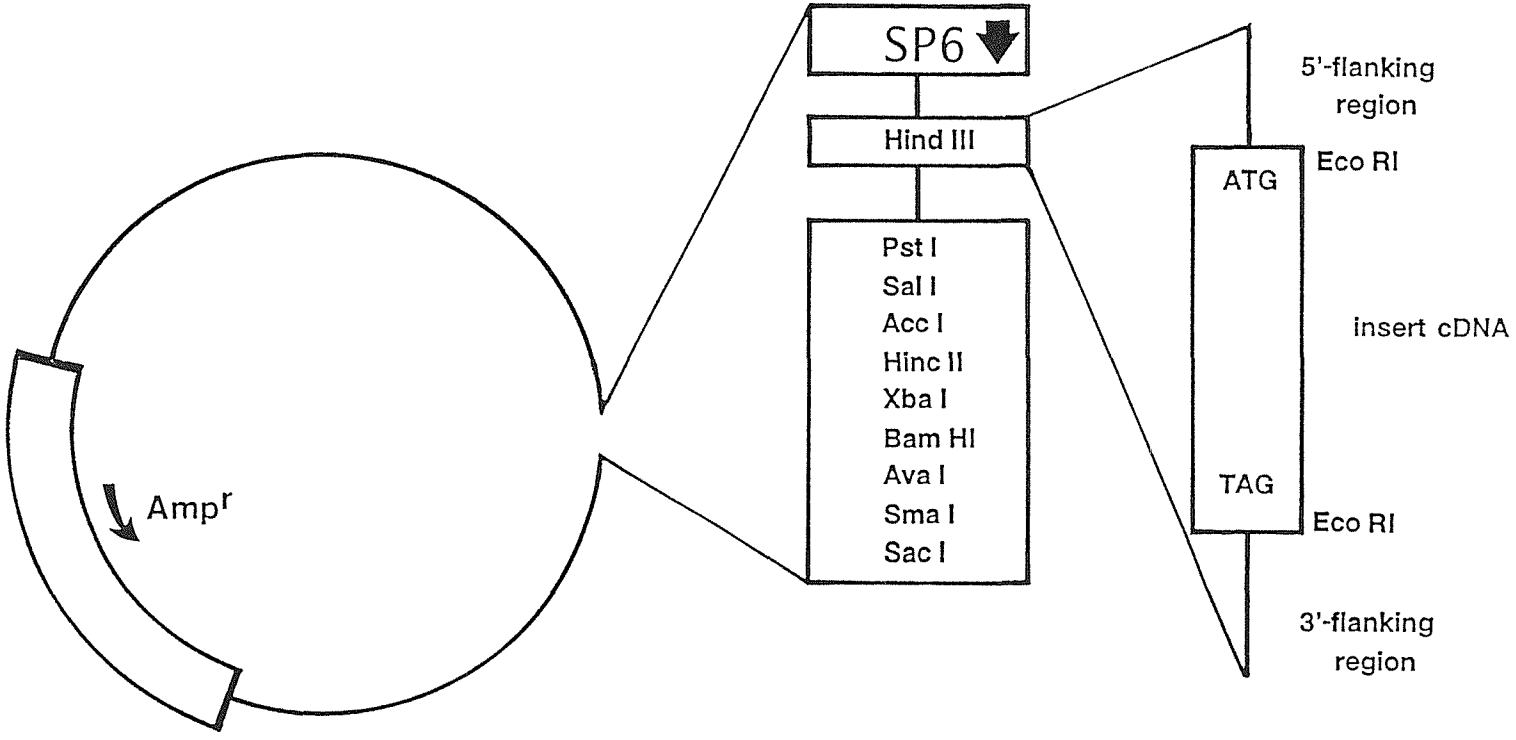
precipitated. The DNA was blunt-end ligated (overnight @ 20°C) using T4 DNA ligase (5U, BRL) and the PEG-containing buffer supplied [composition of 5x stock solution: 250mM tris-HCl, 50mM $MgCl_2$, 5mM ATP, 5mM DTT, 25% (w/v) PEG-8000, pH 7•6]. Transformation-competent TG2 bacteria were transformed with dilutions of the ligation mix and transformed cells were selected on ampicillin plates. Alkaline lysis minipreps were performed on overnight cultures prepared from ampicillin-resistant colonies. Initially clones were screened for the absence of an *Eco*RI site and once this had been established further restriction analysis was performed to confirm the intactness of the rest of the polylinker. Plasmid DNA was prepared from a 10ml overnight culture of a suitable clone. A glycerol stock was also made from this culture (pSP64T*). Plasmid DNA (≃20μg) was purified using Tip-20 columns according to the manufacturer's "high purity for sequencing" protocol. The *Bam*HI / *Eco*RI adaptor (0•2 OD_{260} units, Fig. 11) was treated with T4 polynucleotide kinase (Boehringer) according to Sambrook *et al.,* (1989) in a reaction volume of 20μl. The plasmid DNA was *Bgl* II-digested to completion (as assessed by ethidium bromide-stained agarose gel electrophoresis) and the linearised plasmid DNA was sodium acetate / ethanol precipitated. The plasmid DNA, redissolved in water, was added directly to the kinase reaction mix. Ligation buffer (10μl, PEG-containing buffer, BRL, as above, and 4000U T4 DNA ligase, New England Biolabs, high concentration, 2000U/μl) was added and the ligation was incubated overnight at 20°C. Transformation-competent TG2 cells were transformed with dilutions of the ligation mix and transformed cells were selected on ampicillin plates.

Fig. 10. The plasmid vector pSP64TW.

The plasmid pSP64T (Krieg & Melton, 1984) was modified by removal of the *Eco* RI site from the polylinker and conversion of the *Bgl* II site between the 5'- and 3'-flanking sequences of the *Xenopus* β-globin mRNA to an *Eco* RI site. The remaining restriction sites in the polylinker are used to linearise the vector, prior to *in vitro* transcription.

Fig. 10.

The plasmid pSP64TW



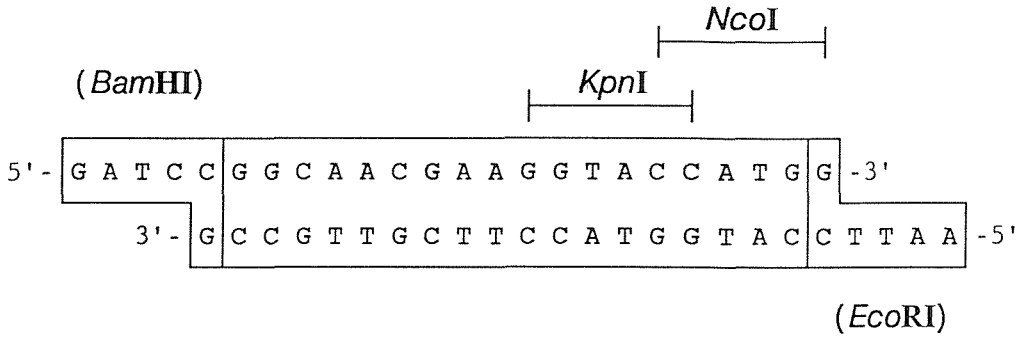


Fig. 11. *BamHI / EcoRI* **adaptor** (Haymerle *et al.*, 1986).

Alkaline lysis minipreps were performed on overnight cultures prepared from ampicillin-resistant colonies. Clones were screened for the presence of an *EcoRI* site. Once this had been established, a large scale preparation of the modified vector was performed as above. Purified inserts (SF6 and α_2-HS-19) were ligated into this vector, pSP64TW, and their orientations established by restriction analysis and direct dideoxy sequencing of the constructs using the SP6 promoter sequencing primer (5'-d[ACCTTATGTATCATACACAT]-3'). Plasmid DNA was purified using Tip-20's and the protocol supplied by the manufacturer. Purified plasmid DNA was sodium hydroxide / EDTA-denatured (Lim & Pène, 1988) and precipitated prior to sequencing.

2.4.2. *In vitro* transcription.

Two protocols were used depending on whether the RNA was required for translation or for the RNAase protection assay.

For Translation. The cDNAs encoding sheep fetuin (clone SF6, see Results) and human α_2-HS glycoprotein (clone α_2-HS-19, from the laboratory of Prof. W. Müller-Esterl, see Fig. 12) were subcloned into the *"translation"* vector pSP64TW *(vide supra),* Fig. 10. The constructs were linearised (pSP64TW-SF6 with *Sma I*, pSP64TW-α_2HS19

with *Pst* I) and transcripts were prepared *in vitro* using SP6 RNA polymerase (BRL) and conditions described by Kassavetis *et al.*, (1982) and Butler & Chamberlin, (1982), except that the cap reagent $m^7G(5')ppp(5')$-G (New England Biolabs) was added to the reaction mix at 0•5mM (Darveau *et al.*, 1985). The SP6 buffer supplied by BRL was used [5x stock solution has composition: 0•2M tris-HCl, 30mM $MgCl_2$, 10mM spermidine, pH 7•9 (@ 25°C)]. The plasmid template was digested with RNAase-free DNAase (Boehringer) and the transcribed mRNA was phenol/chloroform extracted and precipitated (see Krieg & Melton, 1987).

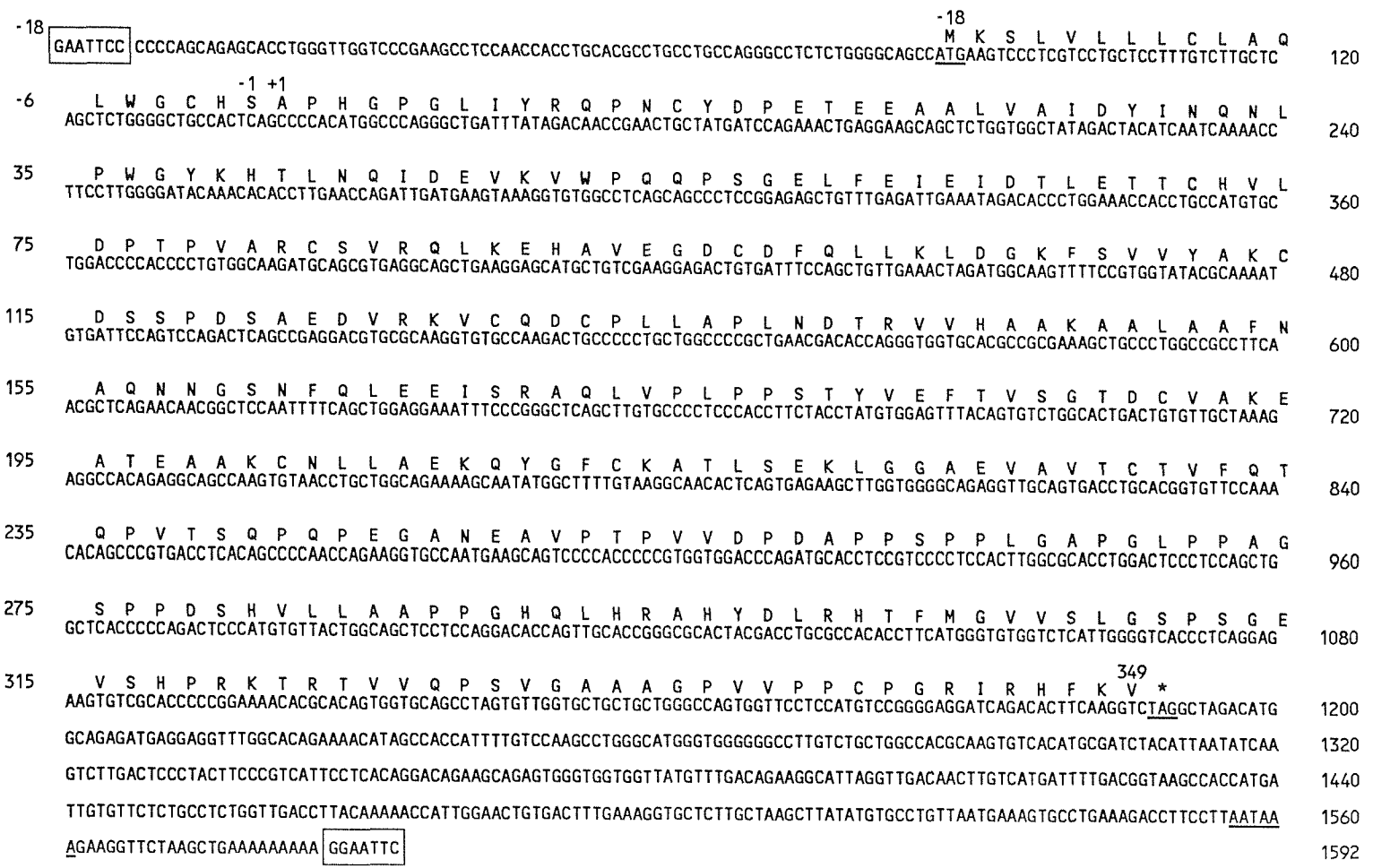
Fig. 12. The nucleotide and deduced amino acid sequence of human α_2-HS glycoprotein.

The complete nucleotide and deduced amino acid sequence of human α_2-HS glycoprotein, clone α_2-HS-19, from the laboratory of Prof. Werner Müller-Esterl, University of Mainz.

Numbers to the left refer to the protein, the putative signal peptide having negative numbers, there being no residue zero. Numbers to the right refer to the nucleotide sequence, and include the *Eco*RI linkers on both ends of the cDNA. The *Eco*RI linkers are boxed. The mature protein comprises 349 residues. The initiator methionine is underlined. The termination codon is indicated by a star (*) and is underlined. The polyadenylation signal AATAAA is underlined.

The sequence differs in two respects from that reported by Lee *et al.*, (1987a); 1) the 5'-UTR is considerably longer and 2) at position 183, this clone has a *T* whereas Lee *et al.*, (1987a) reported *G*, changing amino acid 15 from *asp* (Lee *et al.*, 1987a) to *tyr* (this clone).

Fig. 12. The nucleotide and deduced amino acid sequence of human α_2-HS glycoprotein.



For RNAase protection assays. The *Pst* I/*Bgl* II restriction fragment encoding the carboxyl terminal of sheep fetuin (Fig. 17) was force cloned into *Bam*HI / *Pst* I digested pGEM3Zf(+). A large scale, high purity plasmid preparation was performed. The pGEM3ZF(+)/WMB220 construct (Fig. 13) was linearised with *Bsm*I, *Nae*I, *Pst* I or *Hind* III depending on the experiment in question. A high specific activity antisense transcript was prepared *in vitro* using T7 RNA polymerase (Promega), "SP6 grade" [^{32}P]-α-CTP (Amersham, code PB.20382, specific activity ≃800Ci/mmol) and the conditions set out in Ausubel *et al.,* (1989). After 60 minutes incubation at 37°C, a second aliquot of T7 RNA polymerase was added and the reaction continued for a further 45 minutes. The RNAase protection probe was then gel-purified according to Ausubel *et al.,* (1989), using a 6% nondenaturing polyacrylamide gel. Typically 0•2μg DNA was used and the purified probe had a specific activity of 2-8x10^{8}cpm/μg.

2.4.3. *In vitro* translation.

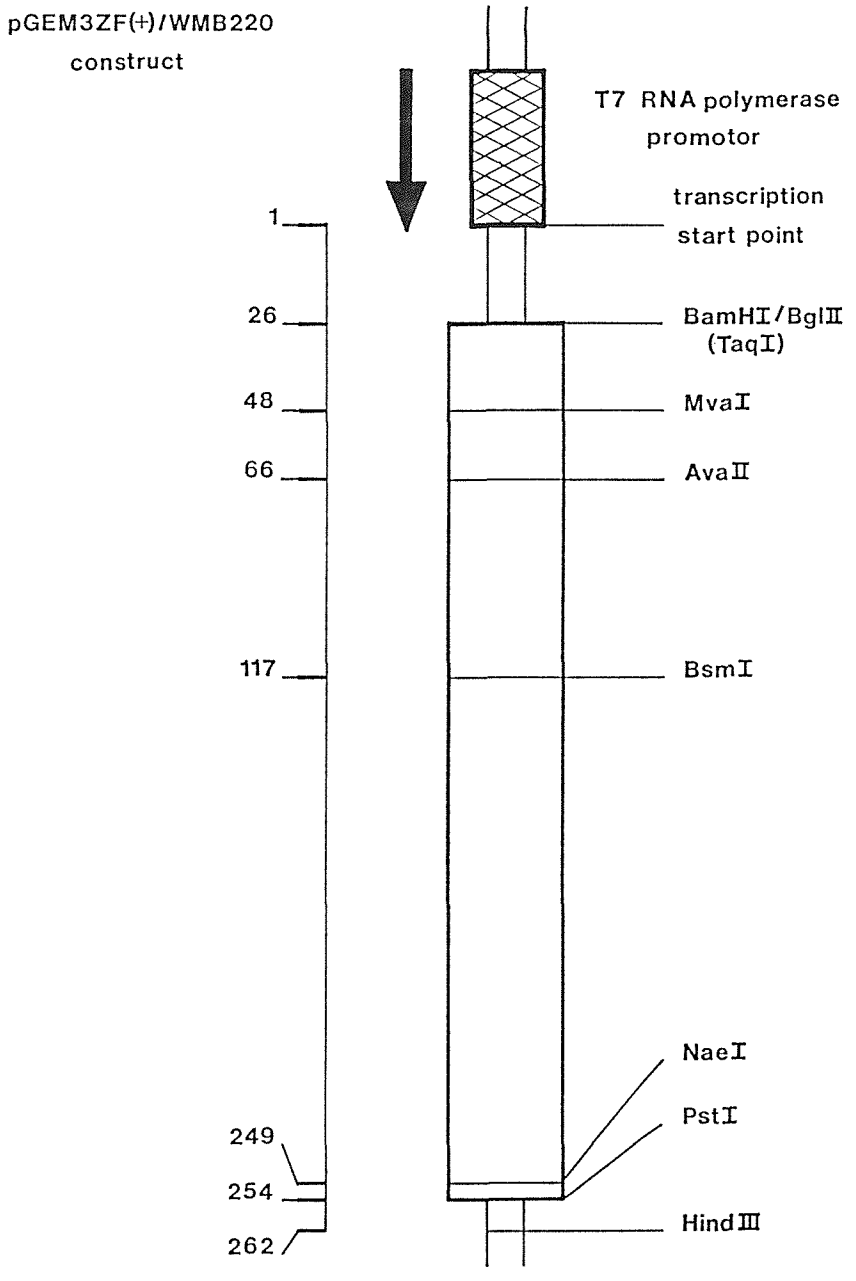
The capped mRNAs so transcribed were translated in a cell-free reticulocyte lysate system (Promega) according to the manufacturer's instructions with the exception that tritiated proline (L-[2,3,4,5-^{3}H] proline, Amersham, code TRK.534, 30μCi in a 20μl reaction) was used as the label and a "minus proline" amino acid mix (Amersham) was substituted for the "minus methionine" mix supplied.

2.4.4. SDS polyacrylamide gel electrophoresis.

Samples were analysed by SDS polyacrylamide electrophoresis using the method of Laemmli, (1976) according to Sambrook *et al.,* (1989). A 3% stacking and a 9% resolving gel were routinely used. When required, gels were fixed and soaked in Amplify® (Amersham) according to the manufacturer's protocol, before being dried under

Fig. 13. The construct pGEM3Zf(+)/WMB220.

The *Pst* I / *Bgl* II fragment of the sheep fetuin cDNA (Figs. 18,19), encoding the carboxyl-terminus of the protein, was force cloned into the pGEM3Zf(+) vector (Promega). An antisense transcript was prepared *in vitro* using T7 RNA polymerase, for details see text.



vacuum onto filter paper. Dried gels were exposed against preflashed Fuji RX film in cassettes with intensifying screens @ -70°C.

2.4.5. RNAase protection experiments.

Using the antisense transcript, synthesised and purified as above, RNAase protection experiments were conducted according to Ausubel *et al.,* (1989). RNAase A and T_1 were purchased from Boehringer, proteinase K (Ultra Pure) was from BRL. Purified probe was used immediately. Hybridisations were performed overnight in a gently shaking waterbath. The formamide used (Fluka, puriss) was deionised using Amberlite MB-1 resin. The amount of RNA in a digest was maintained at 20μg by addition of *E. coli* tRNA (Boehringer). This same tRNA was used as a carrier for precipitations.

2.5. Nomenclature of fetuins.

Proteins from various species characterised by extensive sequence identity with bovine fetuin are denoted "fetuins" throughout. In the case of human fetuin, the original designation of "α_2-HS glycoprotein", identifying the initials of the co-discoverers *H*eremans & *S*chmid (Heremans, 1960; Schmid *et al.,* 1961; Schultze *et al.,* 1962), is used synonymously. The protein encoded by clone pp63 (Auberger *et al.,* 1989) and identified as rat fetuin (Brown *et al.,* 1991a) is also referred to as (rat) pp63.

3. Results.

3.1. Cloning and sequencing of bovine, ovine and porcine fetuins.

In an attempt to define common structural features of mammalian fetuins and to delineate some of their key structural motifs, the complementary DNAs encoding cow, sheep and pig fetuin were cloned, and the deduced amino acid sequences were compared with that known for human α_2-HS glycoprotein, and the other more recently reported members of the fetuin family.

3.1.1. Cloning and sequencing of bovine fetuin.

In order to study fetuin, to establish the molecular basis of the link with human α_2-HS glycoprotein, and to study the origin of fetuin in the developing brain, it was first necessary to clone the cDNA encoding the protein, both to complete the protein sequence and to enable the use of molecular biology techniques. Using the partial bovine fetuin protein sequence (Alcaraz *et al.,* 1981; Johnson & Heath, 1986a; Christie *et al.,* 1987), oligonucleotides were designed (see Fig. 8) and a bovine fetuin cDNA was obtained by screening a commercially available adult bovine liver cDNA library with these probes.

The oligonucleotide probes were labelled using T4 polynucleotide kinase and [^{32}P]-γ-ATP. The labelled probes were spun column-purified and used to screen an adult bovine liver cDNA library in λgt11 (Clontech) at low stringency (42°C, 2xSSC). One positive clone was picked and rescreened until single. Very high background was observed, presumably because of the low hybridisation and washing temperature. This clone (BFG) was analysed by small scale liquid lysate (Sambrook *et al.,* 1989), and insert DNA was isolated by preparative agarose gel electrophoresis. The 1223bp insert was subcloned into the vector M13 mp19 (Yanisch-Perron *et al.,* 1985) and analysed by dideoxy sequencing (Sanger *et al.,* 1977). By comparison with the human α_2-HS glycoprotein cDNA sequence (Lee *et al.,* 1987a) and the known partial bovine fetuin protein sequence (Alcaraz *et al.,* 1981; Christie *et al.,* 1987) it was clear that BFG

was a partial-length bovine fetuin cDNA. The insert remaining from the liquid lysate preparation was random primer-labelled (Feinberg & Vogelstein, 1983) and used to rescreen the same library at higher stringency (55°C, 2xSSC). Twelve strongly positive clones were picked and rescreened until single. These were also analysed by small scale liquid lysate. Clone BF4 contained the largest insert and the full nucleotide sequence of the clone was determined on both strands, according to the sequencing and subcloning strategy shown in Fig. 14. It was found to be a full length bovine fetuin clone and the complete nucleotide and deduced amino acid sequence of the clone is shown in Fig. 15. The sequence, in conjunction with the new protein sequence data determined by Dr. D.L. Christie, reported in Dziegielewska *et al.*, (1990) completes, and in a number of places corrects, the partial protein sequences in the literature (Alcaraz *et al.*, 1981; Johnson & Heath, 1986a; Christie *et al.*, 1987).

Clone BF4 comprises a 5'-untranslated region (UTR) of 12bp including a 5'-GCCACC-3' sequence which constitutes part of the consensus signal for initiation of translation in higher eukaryotes (see Kozak, 1989). This sequence immediately precedes an ATG codon which begins a 54bp region coding for a putative 18 amino acid signal peptide, in complete agreement with the partial sequence reported by Johnson & Heath, (1986a). The open reading frame of 1080bp encodes the entire amino acid sequence of bovine fetuin. Analysis of the putative signal peptide (von Heijne, 1983) and comparison with that of human α_2-HS glycoprotein (Lee *et al.*, 1987a) suggests a cleavage site for the signal peptidase between ser^{-1} and ile^{+1}, which predicts an amino-terminal sequence *H_2 N-ile-pro-leu-asp-pro-val-ala-gly-tyr,* in complete agreement with the previously reported protein sequence (Alcaraz *et al.*, 1981). The 391bp 3'-UTR includes the canonical polyadenylation signal (5'-AATAAA, Proudfoot & Brownlee, 1976) and a 14bp poly(A) tail.

Three sites of N-glycosylation of the form *asn-xaa-ser / thr* were identified in the protein sequence deduced from the cDNA. The carbohydrate moieties on all three of these sites had previously been characterised (see Yet *et al.*, 1988 and Introduction for

references). Sites of O-glycosylation are discussed below. Thus, the predicted mature plasma protein is 341 amino acids long and includes 12 cysteine residues at identical positions to those in human α_2-HS glycoprotein.

Fig. 14. Bovine fetuin sequencing and subcloning strategy. The strategy used to determine the nucleotide sequence of the bovine fetuin cDNA is presented. The sequence was determined fully on both strands. Sequence specific oligonucleotide primers were used where necessary. The cDNA is represented as an open bar; arrows indicate the extent and direction of sequence analysis. The scale indicates the nucleotide number. Key restriction sites are indicated. ATG, initiator codon; TAG, stop codon; AAAAA, poly(A) tail.

Fig. 15. The cDNA and deduced amino acid sequence of bovine fetuin. The complete nucleotide and deduced amino acid sequence of clone BF4 encoding bovine fetuin was determined by the sequencing and subcloning strategy outlined in Fig. 14. The nucleotide sequence of the coding strand and the predicted amino acid sequence is shown. Numbers to the left refer to the protein, the putative signal peptide having negative numbers, there being no residue zero. Numbers to the right refer to the nucleotide sequence, and include the *Eco*RI linkers on both ends of the cDNA. The *Eco*RI linkers are boxed. The mature protein comprises 341 residues. The initiator methionine is underlined. The termination codon (TAG) is indicated by an asterisk (*) and is underlined. The polyadenylation signal, AATAAA, is underlined. N-glycosylation sites *(asn, -81, -138, -158)* are indicated by a diamond (♦). Points of O-glycosylation, *(vide infra)* assigned from the data of Spiro & Bhoyroo, (1974) *(ser-253, ser-263, thr-265)* are indicated by an inverted triangle (▽).
This sequence has been submitted to the EMBL database (accession No. X16577) and is reported in Dziegielewska *et al.*, (1990).

Fig. 14. Bovine fetuin sequencing and subcloning strategy.

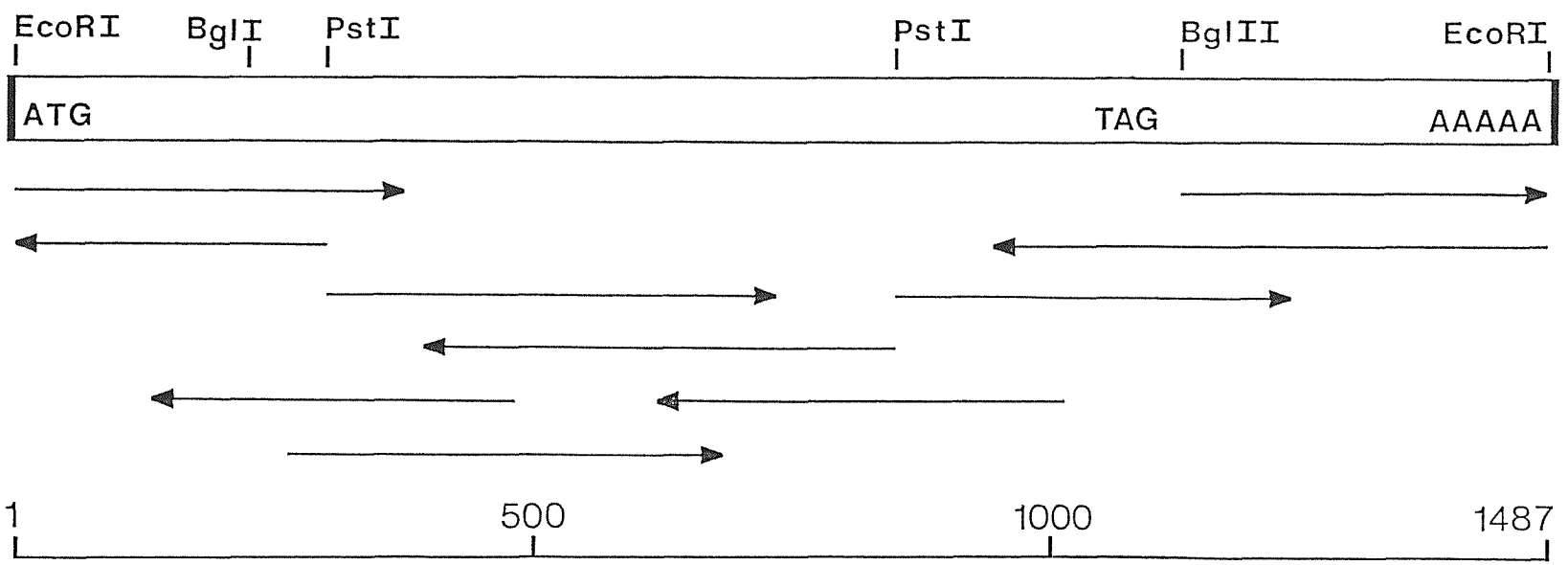


Fig. 15. The nucleotide and deduced amino acid sequence of bovine fetuin.

3.1.2. Cloning and sequencing of ovine fetuin.

Using the bovine fetuin cDNA clone BF4 (Fig. 15) as a probe, a fetal sheep liver library in λgt10 was screened by plaque hybridisation at moderate stringency (55°C, 2xSSC). Twelve strongly positive clones were picked and rescreened using the same probe and conditions until single. Clone SF6, containing the longest insert, was fully sequenced on both strands according to the sequencing and subcloning strategy set out in Fig. 17. Clone SF6 contains a cDNA insert of 1550bp comprising the entire coding sequence of sheep fetuin (Fig. 18). The clone contains a 5'-UTR of 62bp followed by an initiator codon (ATG) which is flanked by a consensus sequence for eukaryotic initiation sites (see Kozak, 1989). The coding part of the cDNA spans 1092bp including a putative 18 amino acid signal peptide. This signal peptide displays strong homology to the signal sequences of α_2-HS glycoprotein (Lee *et al.,* 1987a) and bovine fetuin (Johnson & Heath, 1986a; see Fig. 15) and is responsible for the insertion of the nascent protein into the lumen of the rough endoplasmic reticulum. The reading frame terminates at position 1154 and is followed by a stop codon (TAG), a 3'-UTR of 371bp and a poly(A) tail of 18bp. The 1092bp open reading frame (ORF) predicts a 364 amino acid sequence (including the signal peptide) of sheep fetuin, corresponding to a mass of 38680 Da for the unmodified polypeptide. The canonical polyadenylation signal (AATAAA, Proudfoot & Brownlee, 1976) is present 28 bp upstream of a poly(A) tail.

The predicted signal peptidase cleavage site, by comparison with other fetuin sequences, and by calculation of the "processing probability" (von Heijne, 1983), is between ser^{-1} and ile^{+1} (cys^{-3}-arg^{-2}-ser^{-1} ↓ ile^{+1}), see Table III. This prediction is consistent with the presence of an amino terminal *ile* as previously reported by Marti *et al.,* (1973).

However, as reported in Brown *et al.,* (1991b) amino terminal sequence analysis of the purified plasma protein revealed *two* amino terminal sequences:

A H_2N-*ile-pro-leu-asp-pro-*,
B H_2N-*arg-ser-ile-pro-leu-*.

The finding of a second amino-terminal sequence which extends the other by two residues, suggests a second cleavage site for the signal peptidase in sheep prefetuin, where *ser* is replaced by *cys* in position -1 (see Fig. 16). Such alternative processing of preproteins is very rare (see Adolf *et al.,* 1990) and the possibility that post-translational removal of two amino-terminal residues *(arg-ser)* by an unknown aminopeptidase cannot be entirely ruled out. However, this would imply that the signal peptidase site normally used is the upstream one and the processing probability is markedly less for B than for A (see Table III).

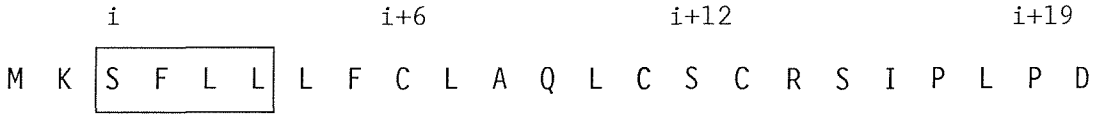
No such alternative amino terminal sequence was found for porcine fetuin (Brown *et al.,* 1991b), nor was it reported for bovine fetuin (Alcaraz *et al.,* 1981; Christie *et al.,* 1987) or human α_2-HS glycoprotein (Yoshioka *et al.,* 1986). Examination of the other mammalian fetuin signal peptides and amino terminal sequences suggests a possible explanation of why only the sheep displays this alternative cleavage site (Fig. 16).

As can be seen in Fig. 16, sheep fetuin can be displaced by two amino acids, and still have "permitted" amino acid residues in the -1 and -3 positions (see von Heijne, 1983). In the other fetuin sequences, the -5 residue is very large (tryptophan) and so this displacement is not possible. Human, bovine and pig fetuins also have a "favourable" glycine at -4, whereas sheep fetuin has serine at -4 (see von Heijne, 1983).

Thus the predicted mature form of sheep fetuin is of 346 (348) residues including 12 cysteine residues and three potential N-linked glycosylation sites of structure *asn-xaa-ser / thr.*

Table III. Processing probability of sheep prefetuin signal peptide cleavage sites. The processing probability was calculated according to von Heijne, (1983). **A** indicates the location of the first four uncharged residues and identifies the frame of the calculation. **B** is a table of the values calculated.

A



B

residues	-5	-4	-3	-2	-1		+1	processing probability	
i+12 - i+17	A	Q	L	C	S	↓	C	8•6	
i+13 - i+18	**Q**	**L**	**C**	**S**	**C**	↓	**R**	**7•6**	"B"
i+14 - i+19	L	C	S	C	R	↓	S	6•3	
i+15 - i+20	**C**	**S**	**C**	**R**	**S**	↓	**I**	**11**	"A"
i+16 - i+21	S	C	R	S	I	↓	P	2•6	
i+17 - i+22	C	R	S	I	P	↓	L	7	
i+18 - i+23	R	S	I	P	L	↓	D	4	
i+19 - i+24	S	I	P	L	D	↓	P	3	

Fig. 16. Fetuin signal peptidase cleavage sites. Putative signal peptides of sheep fetuin (Fig. 18), pig fetuin (Fig. 20), human α_2-HS glycoprotein (Lee *et al.*, 1987a), rat fetuin (pp63, Auberger *et al.*, 1989), mouse fetuin (Yang *et al.*, 1991) and bovine fetuin (Fig. 15) are aligned. "+1" indicates the known or predicted amino terminal residues of the mature secreted protein. "x" indicates an unknown amino acid as the pig cDNA clone was not full length (Fig. 20). The arrow indicates the signal peptidase cleavage site. The -1 and -3 residues are boxed. Single letter code is used.

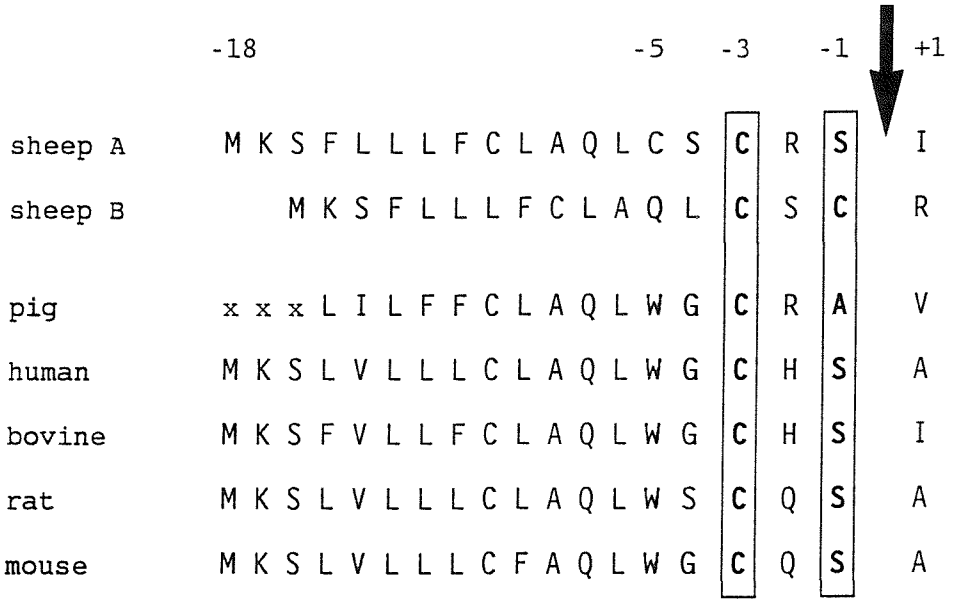


Fig. 17. Strategy for determining the nucleotide sequence of the sheep fetuin cDNA. The clone was sequenced fully on both strands by the subcloning and sequencing strategy indicated. Sequence specific oligonucleotide primers (17mers) were used where required. The insert cDNA of SF6 (sheep fetuin) is represented as an open bar; arrows indicate the extent and direction of sequence analysis. The scale indicates the nucleotide number. ATG, initiator codon; TAG, stop codon; AAAAA, poly(A) tail.

Fig. 18. The cDNA and deduced amino acid sequence of sheep fetuin. The complete nucleotide and deduced amino acid sequence of clone SF6, encoding sheep fetuin was determined by the sequencing and subcloning strategy set out in Fig. 17. The nucleotide sequence of the coding strand and the predicted amino acid sequence are shown. The numbering of the nucleotides (1 to 1550) is given to the right, and the numbering of the corresponding amino acid sequence on the left (1 to 346); the amino acid sequence of the signal peptide is represented by negative numbers in the opposite direction. -1/+1 identifies the "conventional" fetuin signal peptidase cleavage site, though as discussed above and in Brown *et al.*, (1991b), two alternative amino terminal sequences were found on analysis of the purified plasma protein, suggesting two alternative cleavage sites. The *Eco*RI linkers flanking the cDNA insert are boxed. The stop codon (TAG) is underlined and identified by an asterisk (*). The canonical polyadenylation signal (AATAAA) is underlined. Potential N-glycosidic attachment sites of the general structure *asn-xaa-ser / thr* are indicated by diamonds (♦). Potential sites of O-glycosylation are discussed below. This sequence has been submitted to the EMBL database (accession No. X16578) and is reported in Brown *et al.*, (1991b).

Fig. 17. Sheep fetuin sequencing and subcloning strategy.

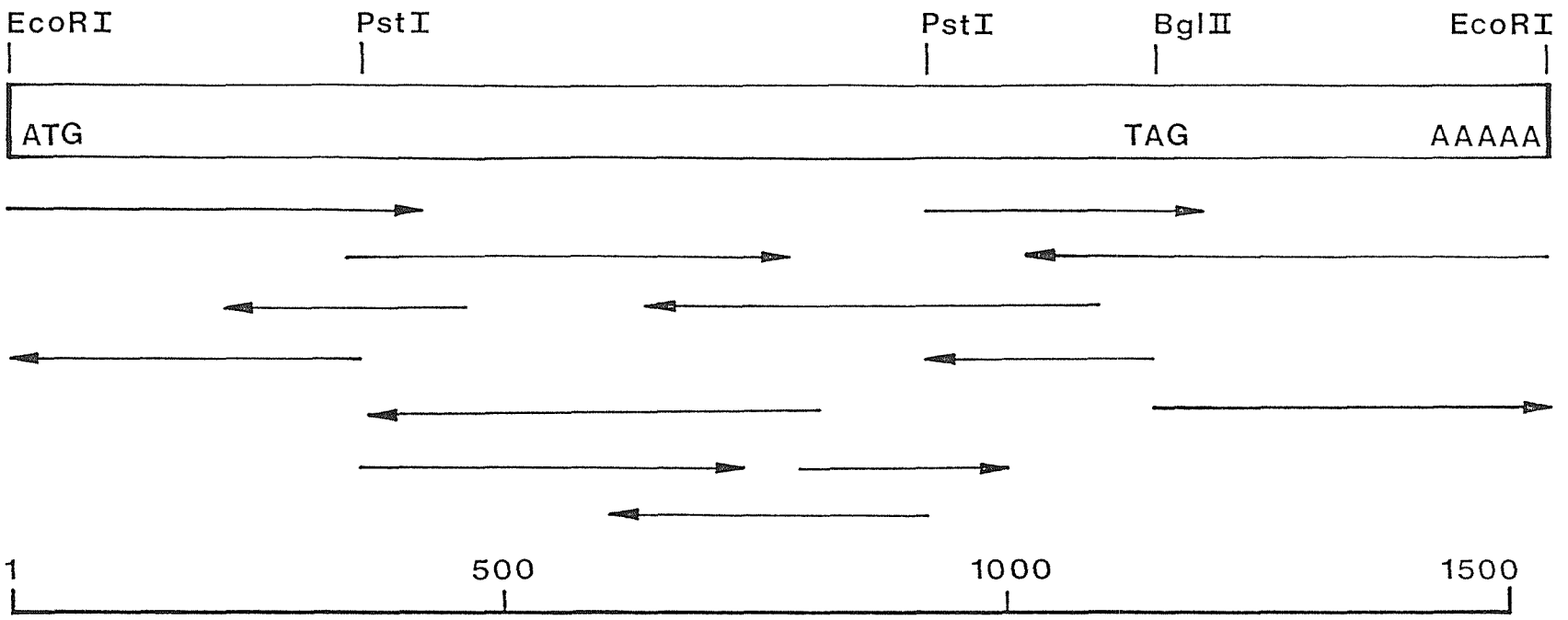
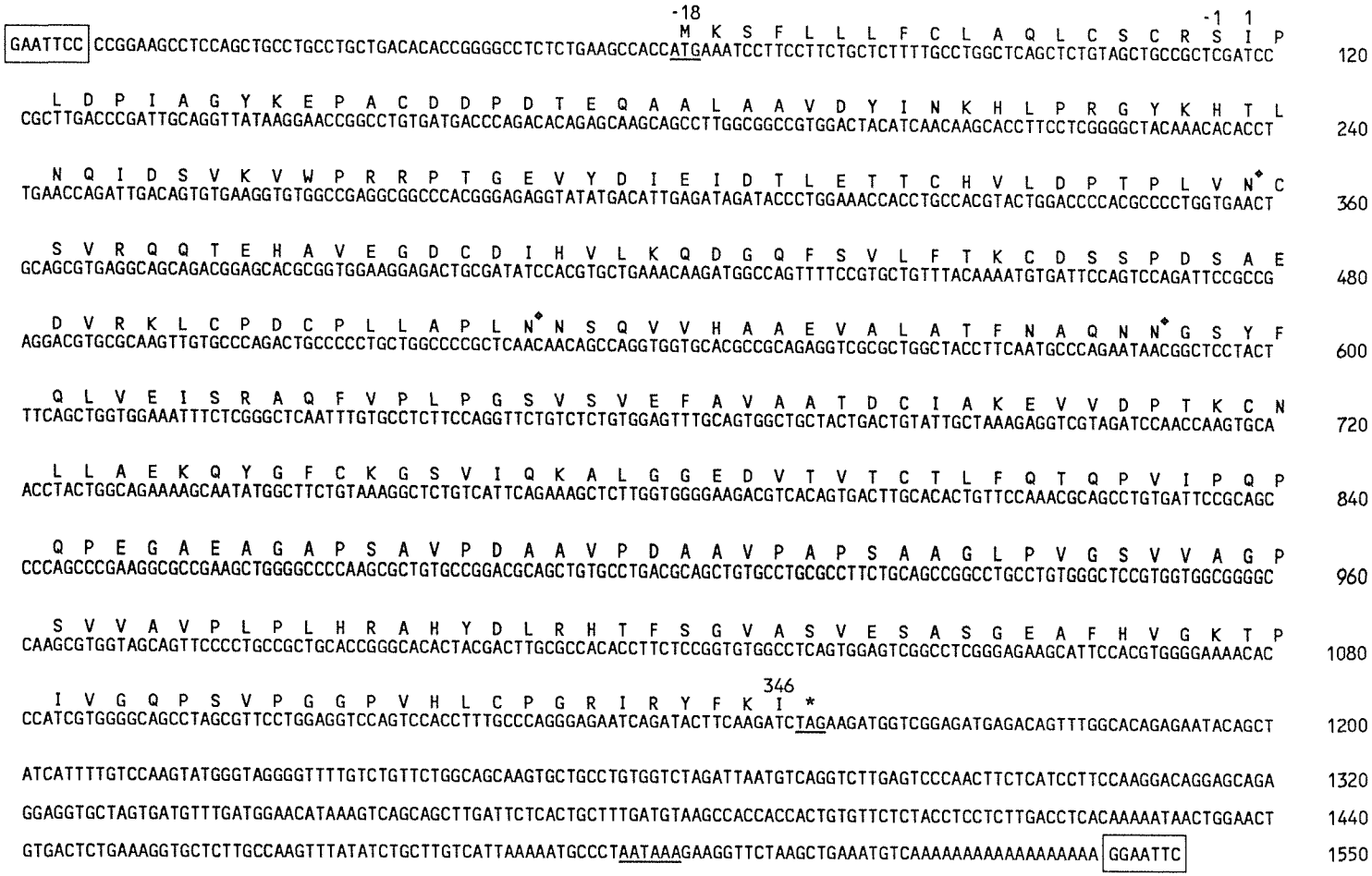


Fig. 18. The nucleotide and deduced amino sequence of sheep fetuin.



3.1.3. Cloning and sequencing of porcine fetuin.

In an analogous manner, an adult pig liver cDNA library in λgt11 (Clontech) was screened under moderate stringency conditions (55°C, 2xSSC) using a probe derived by random-primer labelling the sheep fetuin cDNA, SF6. Among the 12 clones isolated and examined, clone PF5 had the largest insert of 1256bp, but it was not full length (data not shown). A restriction fragment was generated from this clone by *EcoRI/PstI* digestion (see Fig. 19), separated by agarose gel electrophoresis, labelled by random priming and used to rescreen the same library under high stringency conditions (65°C, 1xSSC). Twelve further clones were isolated and examined. Clone PF3 contained an insert cDNA of 1470bp and this was found to be an almost full-length clone for pig fetuin. The ORF of the clone spans 1086bp and terminates at position 1087, there being a stop codon (TAG) and 373bp of 3'-UTR sequence. While a polyadenylation signal (AATAAA) is observed 35bp upstream of the *EcoRI* linker, no poly(A) tail was seen. The ORF encodes 362 amino acids, including 15 amino acid residues of the putative signal peptide (Fig. 20). In the case of human, bovine, mouse and rat fetuin a signal peptide of 18 residues is observed or predicted (Fig. 16). From this it seems that PF3 represents an almost full-length clone, the coding sequence for the initiator methionine and first two amino acids of the signal peptide being missing. The predicted signal peptidase cleavage site (cys^{-3}-arg^{-2}-$ser^{-1}\downarrow$-ile^{+1}), assigned by comparison with other fetuin sequences and calculation of the processing probability (data not shown), was confirmed by the amino terminal protein sequence analysis of the purified mature protein (see Brown *et al.*, 1991b). Thus, the predicted form of pig fetuin is of 347 residues, with 12 cysteines and three potential N-glycosylation sites, in identical positions to those identified in the sheep and bovine sequences.

Fig. 19. Strategy for determining the nucleotide sequence of the pig fetuin cDNA. The clone was sequenced fully on both strands by the subcloning and sequencing strategy indicated. Sequence specific oligonucleotide primers (17mers) were used where required. The insert cDNA of PF3 (pig fetuin) is represented as an open bar; arrows indicate the extent and direction of sequence analysis. The scale indicates the nucleotide number. TAG, stop codon. The *Eco*RI/*Pst* I restriction fragment used for rescreening procedures is indicated. Key restriction sites are indicated.

Fig. 20. The cDNA and deduced amino acid sequence of pig fetuin. The complete nucleotide and deduced amino acid sequence of clone PF3, encoding pig fetuin was determined by the sequencing and subcloning strategy set out in Fig. 19. The nucleotide sequence of the coding strand and the predicted amino acid sequence is shown. The numbering of the nucleotides (1 to 1470) is given to the right, and the numbering of the corresponding amino acid sequence on the left (1 to 347); the amino acid sequence of the signal peptide is represented by negative numbers in the opposite direction. -1/+1 identifies the predicted signal peptidase cleavage site, which was confirmed by amino terminal sequence analysis of the purified plasma protein (see Brown *et al.*, 1991b). The *Eco*RI linkers flanking the cDNA insert are boxed. The stop codon (TAG) is underlined and identified by an asterisk (*). The canonical polyadenylation signal (AATAAA) is underlined.

Potential N-glycosidic attachment sites of the general structure *asn-xaa-ser / thr* are indicated by diamonds (♦). Potential sites of O-glycosylation are discussed below. This sequence has been submitted to the EMBL database (accession No. X56021) and is reported in Brown *et al.*, (1991b).

Fig. 19. Pig fetuin sequencing and subcloning strategy.

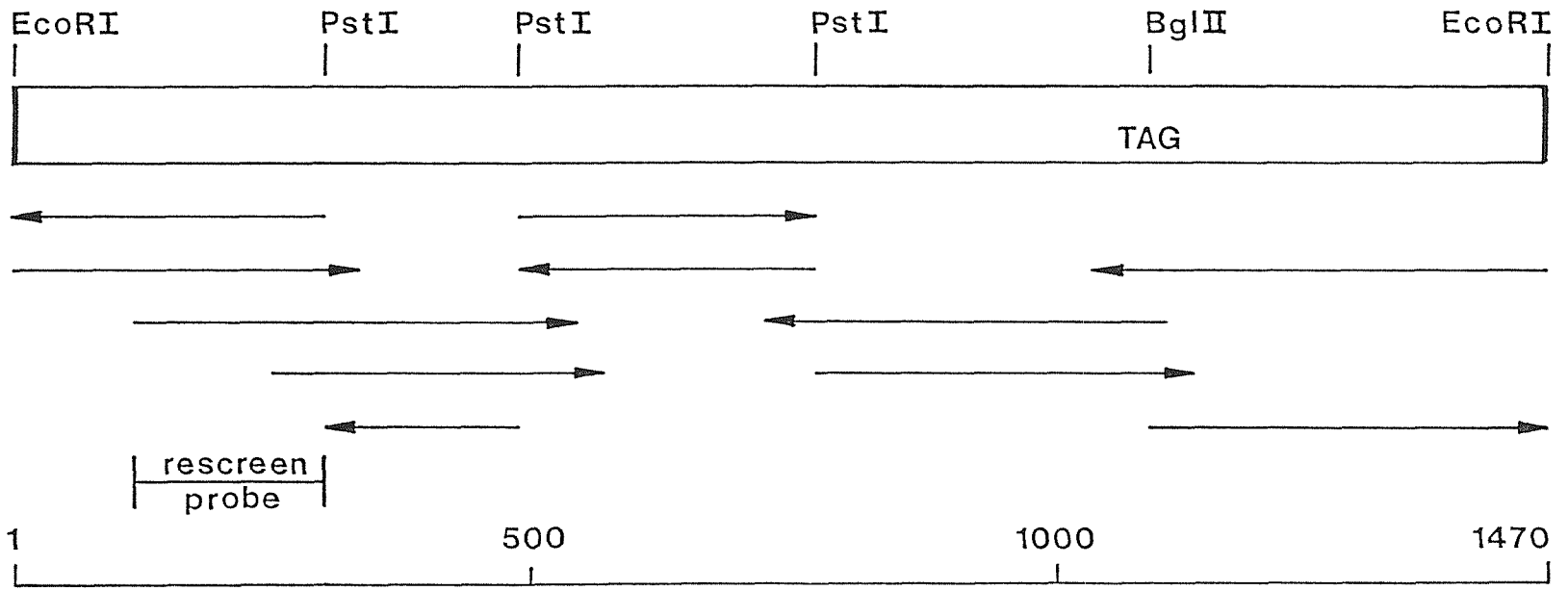
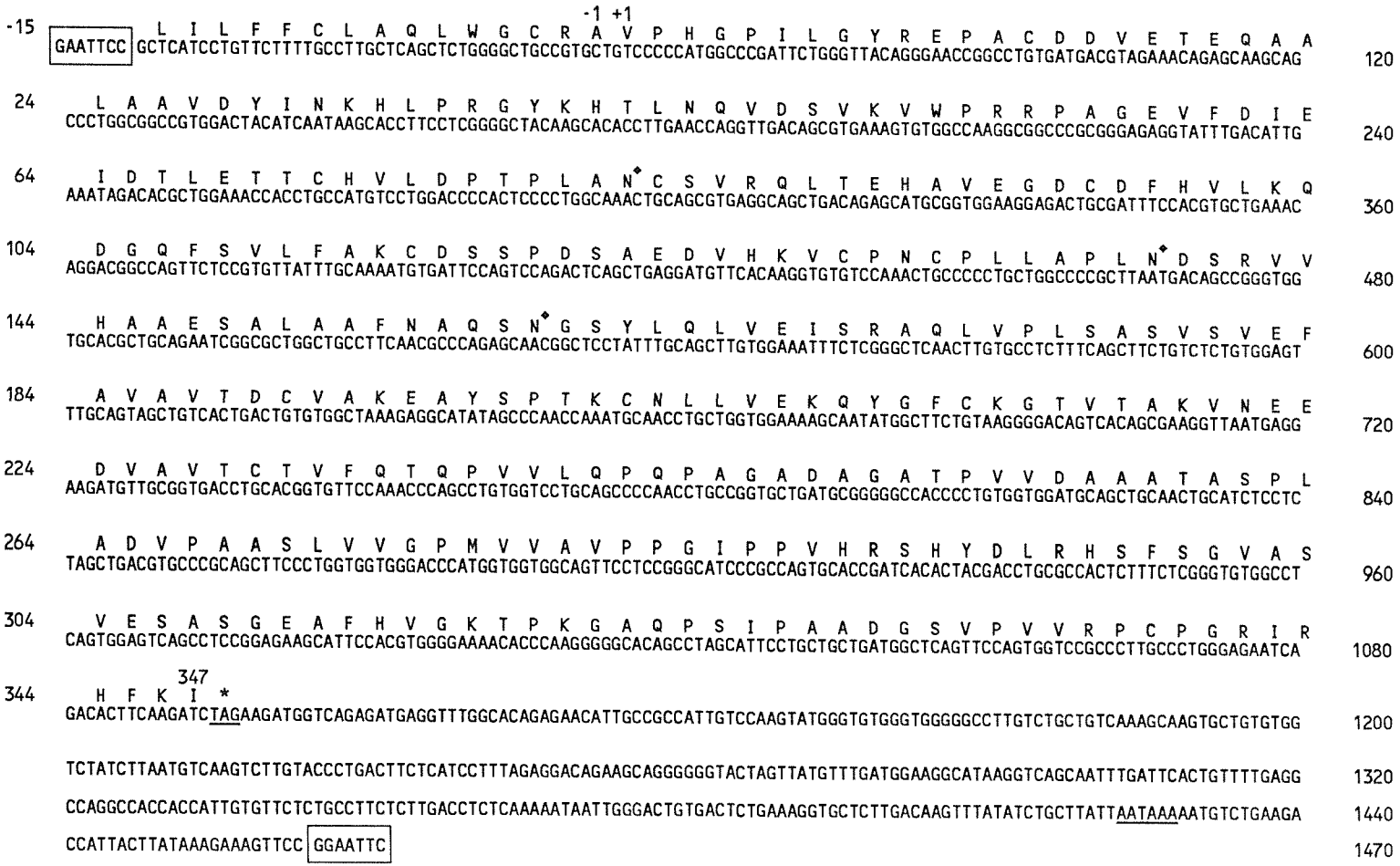


Fig. 20. The nucleotide and deduced amino sequence of pig fetuin.



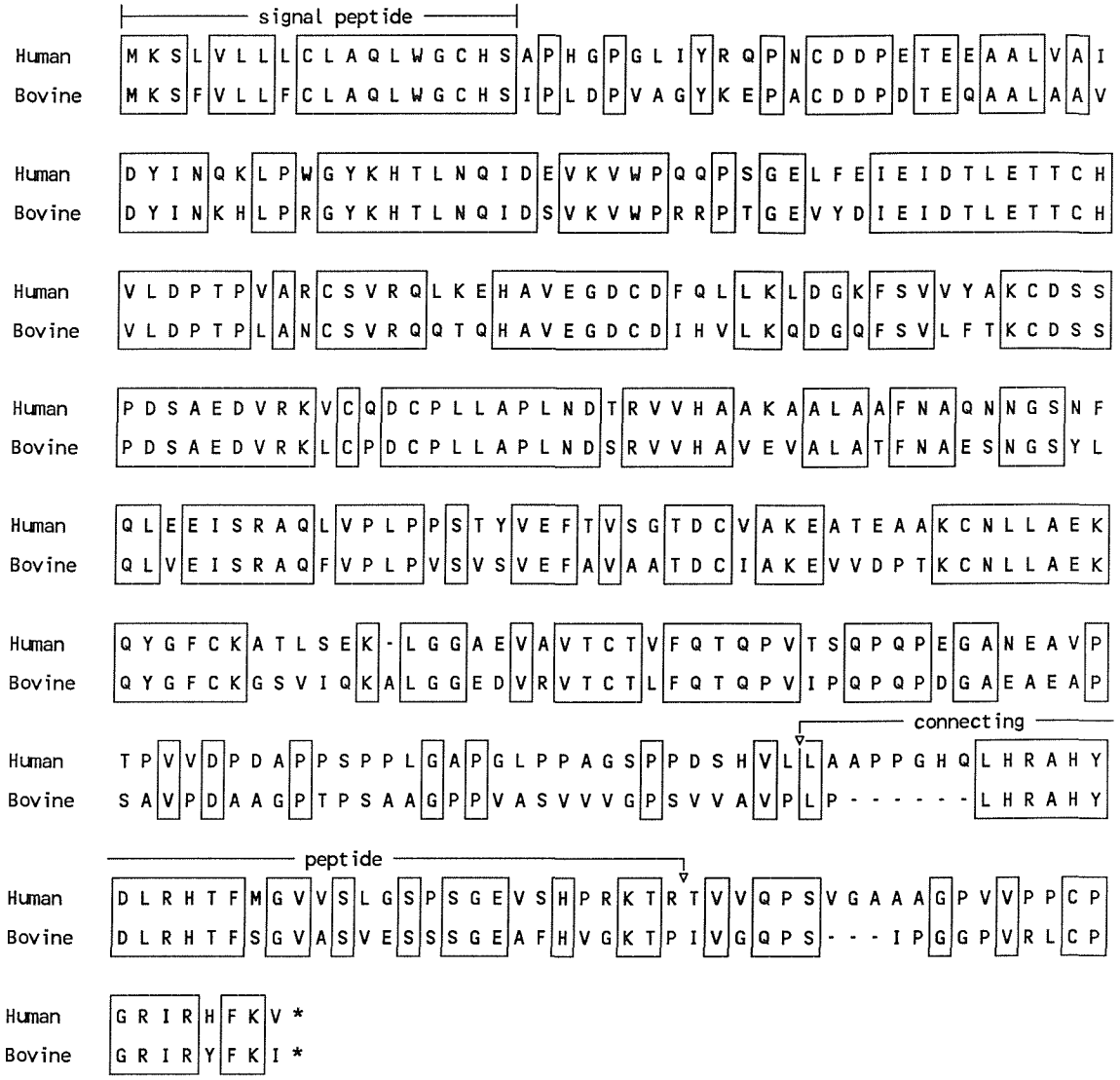
3.2. Southern blot to determine gene copy number of bovine fetuin.

The details of this experiment are described and reported in Dziegielewska *et al.*, (1990). In order to determine the number of fetuin genes in the bovine genome, Southern blot analysis was performed. Bovine genomic DNA (15μg) was digested to completion with *Eco*RI and *Bgl* II. The probe used (see Fig. 14), comprising the whole of the 3'-UTR of bovine fetuin, was labelled by the random primer technique (Feinberg & Vogelstein, 1983) to a specific activity of $\simeq 2 \times 10^9$cpm/μg. The 3'-UTR was used as a probe because it is rarely interrupted by introns which would complicate the result. A single band was obtained when genomic DNA was digested with either enzyme (see Dziegielewska *et al.*, 1990, Fig. 3), implying that bovine fetuin is a single copy gene. No evidence was found for a "bovine α_2-HS glycoprotein gene".

This result, in conjunction with the high degree of sequence identity between the two proteins (Fig. 21, Table V), and the fact that human α_2-HS glycoprotein has been shown also to be a single copy gene (Magnuson *et al.*, 1988), confirms that bovine fetuin and human α_2-HS glycoprotein are species homologues and not simply (closely) related proteins. Furthermore, in the light of this finding it seems highly likely that a number of papers describing α_2-HS glycoprotein in bovine bone (see Dziegielewska *et al.*, 1991 and Introduction for references) were, in fact, identifying fetuin.

Fig. 21. An alignment of bovine fetuin and human α_2-HS glycoprotein.
A protein sequence alignment of human α_2-HS glycoprotein (Lee *et al.*, 1987a) and bovine fetuin (Fig. 15). The signal peptides and, in human α_2-HS glycoprotein, the connecting peptide, are indicated. The carboxyl-terminus is indicated by an asterisk (*). Identical residues are boxed. A minimum number of gaps, indicated by a hyphen (-) was introduced to maximise the alignment.

Fig. 21. An alignment of bovine fetuin and human α_2-HS glycoprotein.



3.3. *In vitro* translation of sheep fetuin & human α_2-HS glycoprotein.

The vector pSP64T (Krieg & Melton, 1984) was successfully modified to create pSP64TW (Fig. 10) as described in Materials & Methods. Both inserts (sheep fetuin, Fig. 18, and α_2-HS glycoprotein, Fig. 12) were subcloned into the modified vector and their orientation was determined by restriction analysis and by direct sequence analysis, using a 17mer primer annealing to the SP6 promoter. Large

scale plasmid preps were carried out and plasmid DNA was purified using pZ523 spun columns. The constructs were linearised (α_2-HS glycoprotein with *Pst* I, sheep fetuin with *Sma* I) and transcripts were prepared *in vitro* using SP6 RNA polymerase. The capped mRNAs so produced were translated in a cell-free reticulocyte lysate system using [^{3}H] proline as the label. The products of this translation were analysed by reducing SDS polyacrylamide gel electrophoresis (data not shown). The anomalous mobility of the reticulocyte lysate translation of bovine fetuin reported by Johnson & Heath, (1986b) applies equally to sheep fetuin and α_2-HS glycoprotein, if not more so. The unmodified molecular weight of all three polypeptide chains is less than 40kDa, yet all three on SDS polyacrylamide gels migrate with an apparent molecular weight of ≃50kDa.

From the cDNA cloning and amino terminal sequence analysis (Brown *et al.,* 1991b), a protein of 346 (348) residues is predicted. The observation of an apparent M_r of 68000 (under reducing conditions, see Brown *et al.,* 1991b) most probably reflects the post-translational modifications of the protein by glycosylation *(vide infra),* although it may also be partly a reflection of this anomalously high apparent molecular weight of the primary translation product.

3.4. Tissue-specific expression of fetuin.

As mentioned in the Introduction one of the aims of this study was to examine whether or not *in situ* synthesis accounts for or contributes to the observed presence of fetuin in the developing forebrain. Fetuin, α_2-HS glycoprotein or a related glycoprotein has been found in the developing cortex of all species so far studied: sheep (Møllgård *et al.,* 1984), rat (Sarantis & Saunders, 1986), cattle (Reynolds *et al.,* 1987), human (Dziegielewska *et al.,* 1987), and tammar wallaby (*Macropus eugenii,* Jones *et al.,* 1988, 1991; Jones, 1990). As discussed in the Introduction, there is evidence for synthesis in the brain of a number of plasma proteins. Three methods were used to examine the expression of fetuin; northern blot, RNAase

protection assays and the polymerase chain reaction (PCR).

3.4.1. Northern blot analysis of sheep fetuin expression.

Total RNA was prepared by the acid-phenol / guanidinium method of Chomczynski & Sacchi, (1987), from fetal (126 day) sheep liver, brain, heart, kidney, and skeletal muscle. Fetal tissue of this age was used simply for reasons of quantity. RNA from 40 and 60 day fetal sheep brain was also examined. The RNAs (20μg of each) were separated by denaturing (2M formaldehyde) agarose gel electrophoresis according to Gerard & Miller, (1986) and capillary transferred to a Hybond N nylon membrane, using 20xSSPE as the transfer buffer. The probe used was the full length sheep fetuin cDNA (Fig. 18) labelled to high specific activity by the random primer technique. Even on prolonged exposure, no fetuin mRNA could be detected in brain (at any of the three ages), heart, kidney or skeletal muscle (data not shown). In the liver a strong signal was obtained. Furthermore, its size was approximately 1600bp, as estimated using the BRL high molecular weight RNA ladder as markers. This suggests that the sheep fetuin clone obtained (Fig. 18) represents almost the entire biological sheep fetuin mRNA. This is in good agreement with the data for the human α_2-HS glycoprotein (Lee *et al.,* 1987a) and the rat fetuin (pp63, Falquerho *et al.,* 1991) mRNAs. On the basis of a primer extension experiment Lee *et al.,* (1987a) concluded that the full length human transcript was 1571 or 1574 nucleotides excluding the poly(A) tail. The intactness of the RNA prepared from all tissues was checked by ethidium bromide staining and by probing a northern blot with an actin probe (gift from Dr. D. Latchman, data not shown).

3.4.2. RNAase protection assays.

Antisense RNA was transcribed *in vitro* using T7 RNA polymerase from the pGEM3ZF(+)/WMB220 construct (see Fig. 13), linearised with *Ava*II, *Bsm*I or *Hind* III. The full-length transcript was gel

purified according to Ausubel *et al.,* (1989) and used in an RNAase protection assay. Total RNA was prepared from fetal sheep liver and brain using the method of Chirgwin *et al.,* (1979). Poly(A)+ mRNA was purified using Poly(A) Qwik columns (Stratagene). RNA (20μg total or 2 to 5μg poly(A)$^+$ mRNA plus *E.coli* tRNA to a total of 20μg RNA) was hybridised overnight at 55°C, in 50% (v/v) deionised formamide and the PIPES buffer described by Ausubel *et al.,* (1989) with ≃2x10^{5}cpm gel-purified probe. The hybridisation mix was diluted approximately 10-fold in RNAase buffer containing RNAase A and RNAase T_1. The RNAases were then proteinase K-digested in the presence of SDS. The RNA was phenol / chloroform extracted, precipitated and analysed by denaturing (8•3M urea) polyacrylamide gel electrophoresis, using a small aliquot of the gel-purified probe as a marker. Liver RNA was serially diluted to estimate the sensitivity of the method; a clearly visible signal was obtained from 10ng fetal sheep (126 day) total RNA. No signal was visible, even on prolonged, exposure from 20μg 40, 60 or 126 day fetal sheep brain RNA. (data not shown). Thus, if fetuin mRNA is present in the (40, 60, 126d) fetal sheep brain at all, then it is present at less than 0•05% of the level in 126 day fetal sheep liver.

3.4.3. PCR experiments.

A series of PCR experiments were undertaken to further examine whether or not fetuin mRNA could be detected in RNA prepared from fetal sheep brain tissue. Total RNA was prepared from 40 and 60 day fetal sheep liver and brain using the method of Chirgwin *et al.,* (1979). Poly(A)+ mRNA was purified using Poly(A) Qwik columns (Stratagene). A first strand cDNA was prepared from RNA by reverse transcription, primed by a sequence-specific antisense oligonucleotide near the termination codon (see Fig. 9). This first strand cDNA was PCR-amplified using the same antisense oligo and a sense oligo 380 bases upstream. The PCR products were subjected to high-stringency Southern blot analysis, using the random-primed sheep fetuin cDNA as a probe. A series of difficulties was encountered using this method, the major problem being that the

only way to "test" the first strand cDNA synthesis, the most problematic step, is to carry the experiment through to the Southern blot, which clearly introduces rather too many further variables. In a number of experiments the product band was apparently of the wrong size and on occasions appeared as a closely spaced doublet (data not shown). No evidence was obtained for there being fetuin mRNA in the fetal sheep (40 or 60 day) brain (data not shown).

The conclusion from these three experiments (northern blot, RNAase protection and PCR) is that if fetuin is expressed in the (40, 60 or 126d) fetal sheep brain at all, then it is at a very low level (less than 0·05% of that in the 126 day fetal sheep liver). This in no way rules out the possibility but does mean that it will be very difficult to study. Examination of the immunocytochemically stained material (see for example Fig. 5) reveals that a very small number of cells are, in fact, fetuin-positive. As the RNA used here was prepared from "total brain" it is possible that the dilution factor is so large that it will not be possible to detect it in RNA from total brain. There is a very obvious way to improve on this, and that is to attempt to dissect the fetuin-positive cells from the brain, to increase their proportion in the tissue from which the RNA is prepared. However, given the fragile state of the fetal sheep brain at this stage, this will, perhaps, not be very practical.

A serious criticism of all three experiments and even attempts to improve on them, is that all three, in a sense, sidestep the main issue, in that none of them shows whether the protein-positive cells contain the mRNA. To answer this question clearly requires *in situ* hybridisation on tissue sections and work is underway within the group to pursue this.

4. Discussion.

4.1. Post-translational modification.

Bovine fetuin has been extensively studied with respect to its carbohydrates (see Introduction). The bovine protein has three O-linked mucin-type carbohydrate groups (Spiro & Bhoyroo, 1974) and three N-linked oligosaccharide structures (Yet *et al.*, 1988). A series of sites for other actual or potential post-translational modifications can be seen in the fetuin sequences.

4.1.1. Proteolytic cleavage.

Three proteolytic cleavage sites have been identified in human α_2-HS glycoprotein. A trypsin-sensitive site was first identified by Lebreton *et al.*, (1979). From more recent analysis (Kellermann *et al.*, 1989), this site has been assigned; it is a typical dibasic processing site *(arg^{125}-lys^{126})* next to the D1/D2 interdomain junction. A corresponding cleavage site was also identified in bovine fetuin (Dziegielewska *et al.*, 1990), and is present in rat fetuin (pp63, Auberger *et al.*, 1989) sheep fetuin (Fig. 18) and mouse fetuin (Yang *et al.*, 1991). In pig fetuin, however, *arg* is replaced by *his* (see Fig. 20). Furthermore, two other proteolytic cleavage sites have been identified in human α_2-HS glycoprotein, flanking the connecting peptide; an unusual site of leu^{282}-leu^{283} and a more conventional site of arg^{322}- thr^{323} (Lee *et al.*, 1987a; Kellermann *et al.*, 1989). This latter sequence is preceded by the homologous dipeptide lys^{320}- thr^{321}. Potential cleavage sites of *lys-thr* are present in all known mammalian fetuins (Fig. 24). The peculiar *leu-leu* site, however, is unique to human α_2-HS glycoprotein, making the other fetuins resistant to proteolysis at this point.

4.1.2. N-glycosylation sites.

Asparagine residues 81, 138 & 158 (in the bovine fetuin sequence) have been assigned as sites for N-linked glycosylation

(Dziegielewska *et al.,* 1990), by comparison of the protein sequence deduced from the cDNA (Fig. 15) and the data of Yet *et al.,* (1988). The first of these (*asn-81* in bovine fetuin) is not present in human α_2-HS glycoprotein, but is present in all other fetuins. The other two sites (138, 158) are conserved in all the fetuins and are known to be N-glycosylated in human α_2-HS glycoprotein (Yoshioka *et al.,* 1986), see Fig. 24, thus suggesting that sheep, pig, rat and mouse fetuins are also glycosylated in these positions. This is further confirmed by recent studies on sheep and pig fetuin, also demonstrating 3 N-linked carbohydrate chains (Y.C. Lee, K.M. Dziegielewska, unpublished).

4.1.3. O-glycosylation sites. O-glycosylation sites cannot be predicted from the amino acid sequence (Jentoft, 1990). However, the presence of *ser* or *thr* in positions known to carry carbohydrate moieties in bovine or human fetuin might indicate positions that are likely to carry O-linked sugars in other fetuins. The position of O-linked sugars in human α_2-HS glycoprotein is known (Gejyo *et al.,* 1983; Yoshioka *et al.,* 1986). The O-linked glycosylation sites of bovine fetuin can be assigned by analysis of glycopeptide composition data presented by Spiro & Bhoyroo, (1974) and the complete amino acid sequence (Fig. 15).

Table IV. The position of O-linked sugars in bovine fetuin.
Analysis of the glycopeptide data reveals the compositions of the peptides (see Table IV, Spiro & Bhoyroo, 1974). Amino acid sequence is in single-letter code.

Glycopeptide	Composition	Assignment in bovine fetuin sequence
B-3	T S P_4 G_2 A_2	260 - GPTPSAAGPP - 269
B-4	T S P_2 G A_2	260 - GPTPSAA - 266
B-5	D S P_2 A_2 V	251 - APSAVPD - 257
B-7	D S E P_2 A_2 V	250 - EAPSAVPD - 257

Spiro & Bhoyroo, (1974) observed that "since B-4 has only 7 amino acid residues and neither serine nor threonine is in the amino terminal position a maximum of 4 amino acids can separate the two [carbohydrate] units". It is now clear that only one amino acid *(pro-263)* separates the threonine and serine residues to which the carbohydrate structures are attached. These results are in keeping with the data presented by Begbie, (1974) who reported that the peptide environment around the O-glycosylation sites of fetuin was particularly proline rich, immediately suggesting the D3 domain. Indeed Begbie, (1974) proposed the sequence *gly-pro-(ser-CHO)-pro-thr-ala* as a site of O-linked glycosylation. It is now clear from an analysis of Spiro & Bhoyroo's data and the amino acid sequence that the correct sequence here is *gly-pro-(thr-CHO)-pro-(ser-CHO)-ala,* as above.

There are a number of substitutions which increase the hydrophobicity and decrease the proline content of bovine, sheep, pig, rat and mouse fetuin compared with human α_2-HS glycoprotein. Indeed, as reported by Dziegielewska *et al.,* (1990), algorithms used to predict membrane-associated helices (Eisenberg, 1984) recognise the region 262 to 285 (bovine sequence) as sufficiently hydrophobic to

be a potential transmembrane domain. Against this, however, the region is rich in proline and glycine which would disrupt a membrane spanning α-helix and, from the above analysis, the region contains three O-linked carbohydrate structures which, presumably, help to maintain its solubility.

Thus, the three O-linked carbohydrate structures in bovine fetuin are located in subdomain D3a (see Fig. 24). While sequence identity between the fetuins is poor in this region (see Fig. 24, Table VII) in the human sequence there is a threonine (*thr* ^{252}) which is known to be O-glycosylated (Yoshioka *et al.*, 1986). In sheep fetuin, serine is present in the same position and threonine in pig fetuin, both amino acids are possible acceptors for O-linked glycosylation.

4.1.4. Proline hydroxylation.

Two potential sites for proline hydroxylation *(ala* ^{267}*-pro-gly* and *pro* ^{286}*-pro-gly)* in agreement with the consensus acceptor sequence of prolyl hydroxylases (McGee *et al.*, 1971), have been identified in human α_2-HS glycoprotein (Kellermann *et al.*, 1989); a single site *(pro* ^{281}*-pro-gly)* is also present in pig fetuin, though not in sheep, bovine, rat or mouse fetuin (see Fig. 24).

4.1.5. Phosphorylation.

In the cystatin superfamily, post-translational modification by phosphorylation has been reported for chicken cystatin (Laber *et al.*, 1989). Examination of the fetuin sequences reveals that clusters of potential phosphorylation sites for a number of protein kinases (Kemp & Pearson, 1990) are present in the well-conserved region of domain D3, positions 301-314, (Fig. 24). Typical recognition sequences of the general structure *ser-xaa-glu* (n=2), *ser/thr-xaa-ser* (n=2), *ser-xaa-xaa-ser* (n=1) & *ser-xaa-xaa-xaa-ser* (n=1) are grouped on a sequence segment of only 14 residues of sheep and pig fetuin, creating a unique series of potential phosphorylation

sites (n=6) for casein kinase-like enzymes (Kemp & Pearson, 1990). The majority of these potential phosphorylation sites are conserved in bovine fetuin (n=5), rat fetuin (pp63, n=5), mouse fetuin (n=5) and human α_2-HS glycoprotein (n=3). A further candidate sequence is located at the interdomain junction of D1 and D2 and harbours a strongly conserved cluster of serine residues flanked by acidic residues (*cys* ^{114} to *ser* ^{120}, see Fig. 24). As reported by Le Cam *et al.,* (1985) phosphorylation of a secreted glycoprotein occurs concomitantly with or immediately after terminal glycosylation in, or subsequent to, the Golgi complex. Note, however, that the casein kinases are cytoplasmic and nuclear enzymes: casein refers to their test, *not* physiological, substrate. Furthermore, it has been reported that neither bovine fetuin (Spiro, 1960) nor human α_2-HS glycoprotein (Schmid & Bürgi, 1961) contains phosphorous.

4.2. Sequence comparisons between the fetuins.

4.2.1. The link between clone pp63 and other fetuins.

Auberger *et al.,* (1989) published the nucleotide and deduced amino acid sequence of a clone *pp63* obtained by immunological screening of a rat liver expression (λgt11) library. The cloned cDNA apparently corresponded to the previously separated and characterised glycoprotein secreted by rat hepatocytes - PP^{63} (Le Cam *et al.,* 1985; Le Cam & Le Cam, 1985). The biological function of the purified protein PP^{63} was reported to be that of a natural inhibitor of the insulin receptor tyrosine kinase (Auberger *et al.,* 1989). Recently, from sequence database searching, it has been reported (Haasemann *et al.,* 1991) that the nucleotide and deduced amino acid sequence presented by Auberger *et al.,* (1989) displays a striking degree of homology with the known sequences of bovine fetuin (Fig. 15) and human α_2-HS glycoprotein (Lee *et al.,* 1987a).

Table V. Sequence identity between the rat protein encoded by clone pp63 (Auberger *et al.*, 1989), bovine fetuin (Fig. 15), and human α_2-HS glycoprotein (Lee *et al.*, 1987a).

protein segment	residue number	identity [%]
overall	358	47
signal peptide	18	73
domain D1	118	53
domain D2	116	59
domain D3	106	27
subdomain D3a	45	16
subdomain D3b	61	36

Data from Haasemann *et al.*, (1991). For details of the domain structure see Elzanowski *et al.*, (1988) and below.

The sequence homology is strong throughout the length of the proteins, except in subdomain D3a and at the extreme carboxyl-terminus. It has previously been reported that the most variable region between members of the fetuin family lies towards the carboxyl-termini of the proteins (Dziegielewska *et al.*, 1990), see Fig. 24. Recently the cDNA sequence of mouse fetuin has been determined (Yang *et al.*, 1991), and a sequence alignment between the D3 domain of this and the protein encoded by clone pp63 (see Fig. 22) reveals that the sequence identity within D3 is 68%, compared with 51% in the same region between human α_2-HS glycoprotein and bovine fetuin (see Fig. 21).

Furthermore, it has been suggested that the discrepancy between the protein sequence deduced from the clone pp63 and other fetuins at the extreme carboxyl-terminus can be explained by a two base duplication in the sequence (Brown *et al.*, 1991a). As can be seen in Fig. 23A, at position 1114-1117 of rat pp63, there is a repeat of the dinucleotide TG - TGTG. By comparison with other fetuin cDNA sequences, it seems possible that this is a cloning or sequencing

artefact. Indeed if there were only one TG, then the reading frame would change and the rest of the deduced amino acid sequence would fall into line with that seen in the five other known fetuin carboxyl-terminal sequences (Fig. 23B).

Overall, sequence identity between mouse fetuin and the rat protein deduced from clone pp63 is so strong (84%) as to leave little doubt that these two proteins are species homologues, and this idea is confirmed by the fact that the mouse and rat, and, indeed, the human and bovine genes have all been shown to be single-copy genes in their respective genomes (see Magnuson *et al.,* 1988; Yang *et al.,* 1991; Dziegielewska *et al.,* 1990; Auberger *et al.,* 1989). Thus, clone pp63 (Auberger *et al.,* 1989) encodes rat fetuin.

Fig. 22. A comparison of the protein sequence of the D3 domain of mouse fetuin (Yang *et al.*, 1991) and the protein sequence deduced from clone pp63 (Auberger *et al.*, 1989), modified from Brown *et al.*, (1991a).

Identical residues are boxed. A minimum number of gaps, indicated by a hyphen (-), was introduced to maximise the alignment. The numbers refer to the longer pp63 sequence, assuming the protein is initiated from the second AUG codon *(vide infra)*, and has a signal peptide of 18 amino acids. The carboxyl-terminus is indicated by an asterisk (*). Sequence is in the single-letter code.

Fig. 22. Sequence alignment of the D3 domains of pp63 & mouse fetuin.

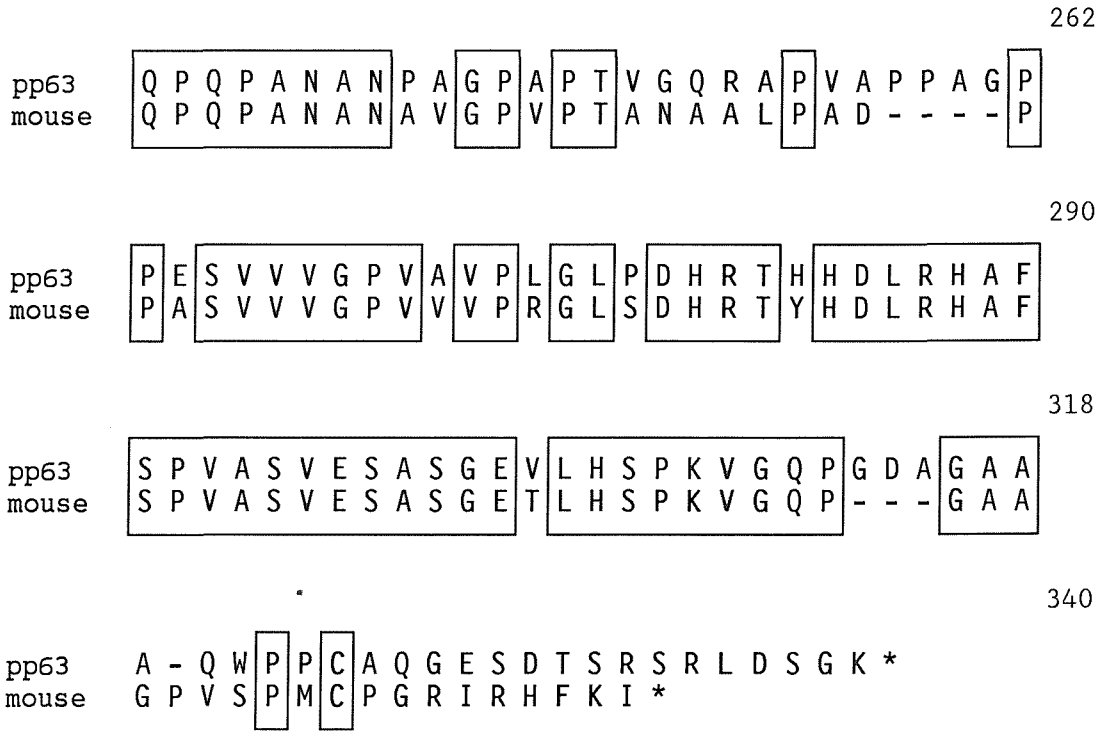


Fig. 23. The discrepancy at the carboxyl-terminus of the protein sequence deduced from clone pp63.

A) "pp63" is the nucleotide sequence encoding the carboxyl-terminus of the protein. The TG repeat - TGTG, nucleotides 1114-1117, is boxed. "pp63-TG" is the "pp63" sequence with only one TG. In both cases the protein sequence encoded is indicated above the nucleotide sequence in lower case letters. The termination codon in each reading frame is underlined and indicated by an asterisk (*).
B) A sequence comparison of the carboxyl-terminal region of the five known fetuins and the translation of "pp63-TG". Identical residues are boxed. The carboxyl-terminus is indicated by an asterisk (*). Sequence is in the single-letter code.

A

pp63

```
   p  p     c  a  q  g  e  s  d  t  s  r  s  r  l  d  s  g  k  *
CCCCCC TG TG CCCAGGGAGAGTCAGATACTTCAAGATCTAGGCTTGACTCAGGGAAGTAA
```

"pp63-TG"

```
   p  p     c  p  g  r  v  r  h  f  k  i  *
CCCCCC   TG CCCAGGGAGAGTCAGATACTTCAAGATCTAGGCTTGACTCAGGGAAGTAA
```

B

"pp63-TG"	p	p	c	p	g	r	v	r	h	f	k	i	*
mouse	p	m	c	p	g	r	i	r	h	f	k	i	*
human α_2-HS	p	p	c	p	g	r	i	r	h	f	k	v	*
bovine	r	l	c	p	g	r	i	r	y	f	k	i	*
pig	r	p	c	p	g	r	i	r	h	f	k	i	*
sheep	h	l	c	p	g	r	i	r	y	f	k	i	*

The question remains, however, as to whether the clone pp63 truly encodes the previously characterised protein PP^{63}, as there is a series of inconsistencies between the papers describing the protein and the clone (see Brown *et al.*, 1991a). Perhaps the most damning of these is that in the original papers describing the purification and properties of the phosphoprotein PP^{63}, the identification of the protein secreted from the hepatocytes was performed using [^{3}H] fucose, [^{32}P] orthophosphate and [^{35}S] methionine. All three isotopes labelled the final protein product in the incubation medium (Le Cam & Le Cam, 1985; Le Cam *et al.*, 1985). However, the amino acid sequence deduced from the pp63 cDNA clone contains *no* methionine (Auberger *et al.*, 1989), after removal of the signal peptide. Sheep fetuin (Fig. 18) and bovine fetuin (Fig. 15) also contain no methionine. Mouse fetuin, on the other hand, contains 4 methionine residues (Yang *et al.*, 1991). This discrepancy could be explained if the protein was initiated from the first and not the second AUG codon in the cDNA; the nucleotide sequence reported by Auberger *et al.*, (1989) encodes two methionines at its 5' end, in the same reading frame. However, as pointed out by Auberger *et al.*, (1989) and discussed elsewhere (Haasemann *et al.*, 1991; Brown *et al.*, 1991a) it is most likely that translation is initiated from the second methionine codon, as this is flanked by a consensus sequence for eukaryotic initiation sites (see Kozak, 1989). The upstream AUG codon (the -51 AUG) has no such flanking sequence. In support of this, the -51 AUG is not seen in the 5' UTR of the cDNAs encoding sheep fetuin (see Fig. 18), human α_2-HS glycoprotein (Lee, 1990, and clone α_2-HS-19, see Fig. 12) or mouse fetuin (Yang *et al.*, 1991). Furthermore, initiation from the second methionine would result in a signal peptide of striking homology to those known in other members of the fetuin family (see Brown *et al.*, 1991a, and Fig. 16) and of similar "processing probability" (von Heijne, 1983).

More recently a later paper from the same group (Falquerho *et al.*, 1991) "corrected" the cDNA sequence, having since discovered that the 5' end of the cDNA (the first 140bp) was not present in the pp63 gene, or in a "new" cDNA. Thus, there is no longer any doubt that translation begins from what was the second methionine.

Fig. 24. Protein sequence alignment of the members of the fetuin family.

The protein sequences of bovine, sheep and pig fetuin (Figs. 15, 17, 20), human α_2-HS glycoprotein (Lee *et al.*, 1987a), mouse fetuin (Yang *et al.*, 1991) and rat fetuin (pp63, Auberger *et al.*, 1989) are aligned.

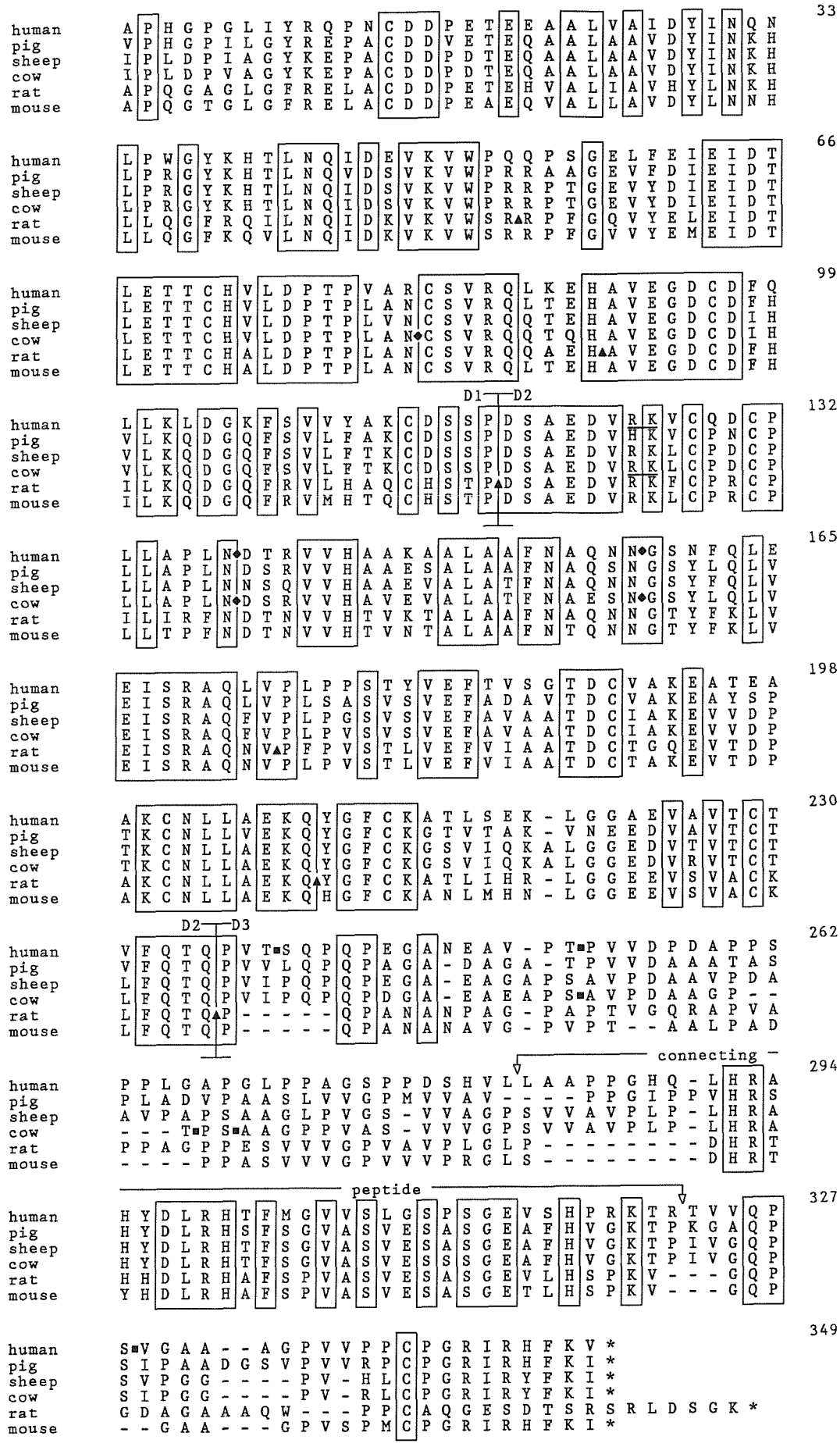
Note that the signal sequences have been omitted and are presented and discussed in Fig. 16. The striking difference between the rat sequence and the other fetuins at the extreme carboxyl-terminus is discussed above (Fig. 23) and in Brown *et al.*, (1991a).

Numbers refer to the human α_2-HS glycoprotein amino acid sequence deduced from the cDNA, *i.e.* including the connecting peptide. A minimum number of gaps, indicated by a hyphen (-), has been introduced to maximise the alignment. Identical residues are boxed and the proposed interdomain junctions are indicated by a vertical line. The α_2-HS glycoprotein sequence shown here is the translation of the cDNA (Lee *et al.*, 1987a) and includes the connecting peptide, the position and extent of which is indicated.

The dibasic cleavage site characterised in human α_2-HS glycoprotein (Kellermann *et al.*, 1989) and bovine fetuin (Dziegielewska *et al.*, 1990) is indicated by underlining.

The diamonds (♦) indicate asparagine residues known to be N-glycosylated in bovine and human fetuin. The squares (■) indicate serine and threonine residues known to carry O-linked sugar moieties in bovine and human fetuin. In the case of bovine fetuin, the O-linked glycosylation sites were assigned by analysis of glycopeptide composition data (Spiro & Bhoyroo, 1974) as discussed below.

Filled triangles (▲) indicate the position of introns in the rat fetuin (pp63) gene (Falquerho *et al.*, 1991).


human
pig
sheep
cow
rat
mouse
D1
D2
D3
connecting
peptide

4.2.2. Pairwise sequence comparisons of the mammalian fetuins.

Comparison between pairs of the amino acid sequences of the fetuins reveals strong identity in each case. According to Dayhoff *et al.*, (1978) members of a *superfamily* can frequently be subdivided into *families* of strongly related proteins, and the working definition of this is that members of a family should share approximately 50% protein sequence identity. By this criterion the fetuins constitute a new family within the cystatin superfamily. See Table VI.

4.3. Domain organisation of the fetuins.

The sequences and sequence comparisons reported here indicate that proteins of the mammalian fetuin family comprise three major domains (Fig. 25). From the amino- to the carboxyl-terminus these are; a cystatin-like domain (D1) containing a potential calcium binding site *(vide infra),* a second homologous cystatin domain (D2) containing a dibasic trypsin-sensitive site *(vide supra)* and a terminal domain of complex structure (D3) of largely hydrophobic nature, containing a proline-rich region and the connecting peptide region. Three lines of evidence from the protein sequences and protein chemistry support this proposed tripartite structure: 1) patterns of sequence identity congruent with the proposed domain structure, 2) the presence of homologous domains in other mammalian plasma proteins (HRG, Koide *et al.*, 1986; Koide & Odani, 1987, and the kininogens, see Fig. 26 and Elzanowski *et al.*, 1988), and 3) proteolytic cleavage sites at or near the proposed interdomain junctions *(vide supra).*

Sequence comparisons between the mammalian fetuins (Fig. 24, Table VII) identify clusters of sequence identity unevenly distributed over the length of the molecules, in agreement with the three domains, first proposed for human α_2-HS glycoprotein by Elzanowski *et al.*, (1988). The two cystatin domains are characterised by extensive sequence identity among the mammalian fetuins; 49% of the residues are in identical positions. Eleven of the twelve invariant

Table VI. Pairwise sequence comparisons.

The sequence identity (in % of identical residues) on the protein level (above the diagonal) and nucleotide level (below the diagonal). The calculations on the protein level are based on the alignment in Fig. 24; only the sequences of the coding regions are used for calculations on the nucleotide level. Sequence identity is calculated as the number of residues or nucleotides in identical positions, expressed as a percentage of the shorter sequence. Amino acid residues refers to the entire protein, *i.e.* including the signal peptide and, in human α_2-HS glycoprotein, the connecting peptide.

Table VI. Sequence identity among the fetuin family.

	sheep	**pig**	**human**	**bovine**	**rat**	amino acids
sheep		75•3	64•0	94•2	59•9	364
pig	80•1		65•1	75•1	55•4	362
human	73•8	75•6		64•3	56•3	367
bovine	96•6	81•0	74•9		59•6	359
rat	73•2	70•9	72•5	74•9		359
nucleotides of coding region	1092	1086	1101	1077	1077	

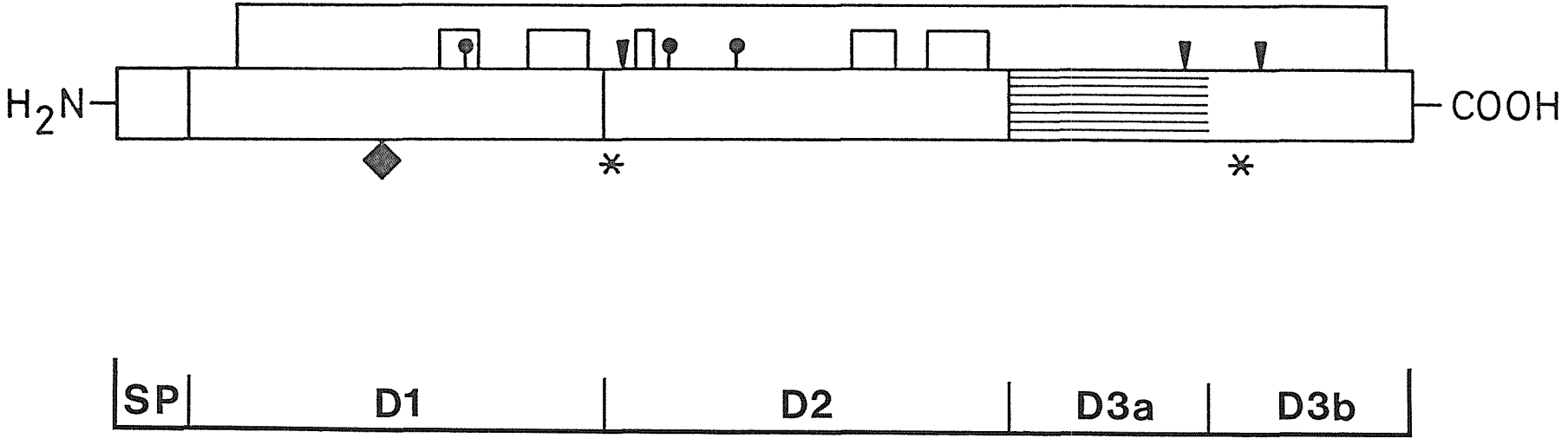
cysteine residues fall within domains D1 and D2. Furthermore the two domains exhibit considerable mutual sequence homology and significant sequence similarity with the cystatin domains in other members of the cystatin superfamily (*vide infra* and Fig. 26).

Table VII. Sequence identity within the fetuin domains.
The sequence identities of the proposed domains of the six mammalian fetuins (human, sheep, bovine, pig, rat, mouse) are given as a percentage of invariant residues, normalised for the shortest sequence. The calculations are based on the alignment in Fig. 24.

protein segment	residue number	identity [%]
overall	345-367	48
signal peptide	15-18	40
domain D1	118	49
domain D2	116-117	49
domain D3	93-115	24
subdomain D3a	41-57	10
subdomain D3b	52-62	35

Fig. 25. Proposed domain structure of mammalian fetuins.
A domain model of fetuin structure is presented. The proposed domains are designated D1 through D3; the signal peptide is marked "SP". The proline-rich region of D3 (also termed D3a) is shaded. The disulphide loop pattern identified in human α_2-HS glycoprotein (Araki *et al.*, 1989, Kellermann *et al.*, 1989) is drawn on top. The N-glycosylation sites (Y), the proteolytic processing sites (▾) present in human α_2-HS glycoprotein and/or bovine fetuin, the potential Ca^{2+} binding site (♦) and putative phosphorylation sites (*) are also indicated. Modified from Brown *et al.*, (1991b).

Fig. 25. Domain structures of the mammalian fetuin.



4.4. Fetuin's membership of the cystatin superfamily.

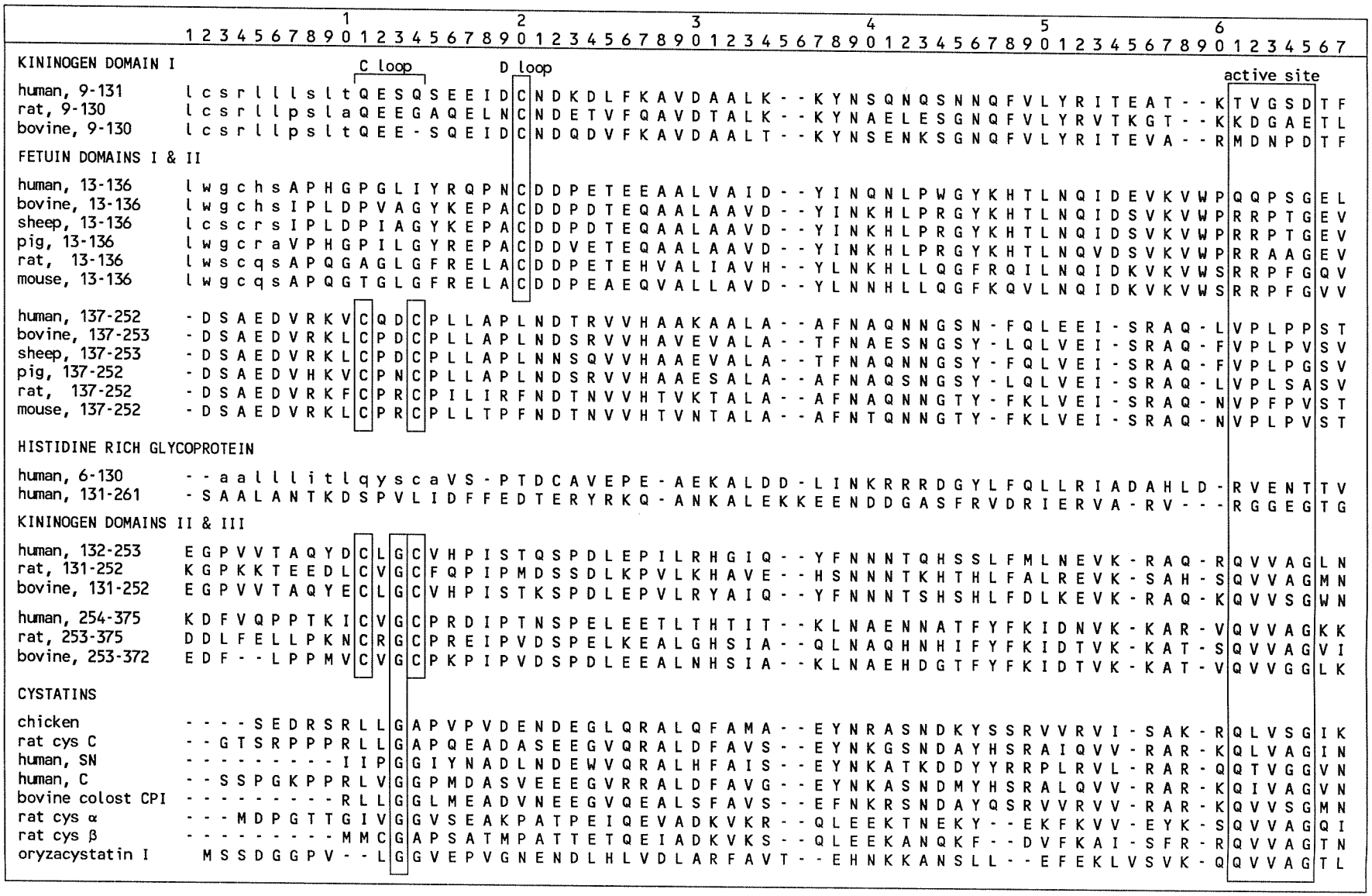
The cystatins are a group of low molecular weight cysteine protease inhibitors with homology to chicken cystatin; the name was first suggested for a protein purified from chicken egg white that was shown to inhibit cysteine proteases of the papain type (Barrett, 1981). The idea of a cystatin *superfamily* was born when it was first realised that multiple copies of a cystatin-like sequence were present in more complex molecules; in the first instance, the kininogens (Nawa *et al.*, 1983; Ohkubo *et al.*, 1984). Nawa *et al.*, (1983) and Ohkubo *et al.*, (1984) observed an apparent sequence duplication within the bovine & human kininogen protein sequences, respectively. Furthermore, both noted that the repeats identified showed homology to a number of cystatin sequences. Further sequence analysis led Salvesen *et al.*, (1986) to propose that there were, in fact, *three* cystatin-like repeats within the kininogen sequence, and this has subsequently been shown to be correct, from the kininogen genomic organisation (see Nakanishi *et al.*, 1987).

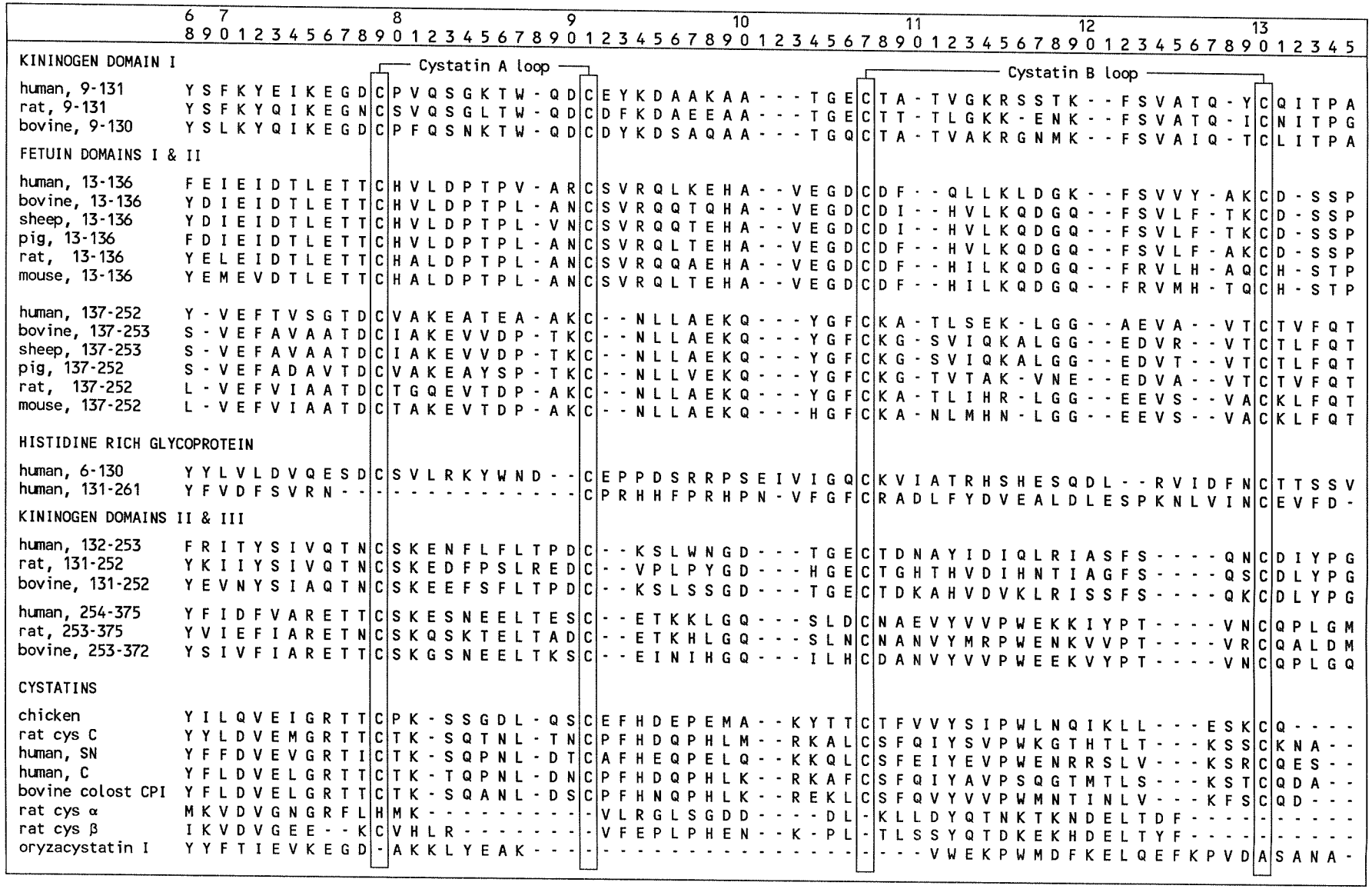
More recently, computer database searching led to the discovery of cystatin-like domains in histidine rich glycoprotein (Koide *et al.*, 1986; Koide & Odani, 1987) and human α_2-HS glycoprotein and bovine fetuin (Elzanowski *et al.*, 1988). Elzanowski *et al.*, (1988), rather belatedly, pointed out that fetuin and human α_2-HS glycoprotein were "*closely related*" on the basis of the amino-terminal sequence of bovine fetuin published by Alcaraz *et al.*, (1981). What Elzanowski *et al.*, (1988) completely failed to notice was that seven months earlier the same claim had been made, and that, further, it had been substantiated with a considerable amount of new bovine fetuin protein sequence (Christie *et al.*, 1987), and that cystatin-like sequences had been found in another protein (Koide & Odani, 1987). A link between the kininogens and human α_2-HS glycoprotein had been first suggested by Hamberg *et al.*, (1975).

Fig. 26. A sequence alignment of cystatin superfamily members.
The cystatin-like domains of human, bovine and rat kininogens, the six mammalian fetuins and human histidine-rich glycoprotein are aligned with selected cystatins. A minimum number of gaps, indicated by a hyphen, has been introduced to maximise the alignment. Numbers given in the left-hand column include the signal peptide (*i.e.* initiator *met* = 1). Where the signal peptide is included it is indicated by lower case letters. The single letter code is used. This table is based on those presented by Barrett *et al.,* (1987) & Elzanowski *et al.,* (1988) and includes much sequence published since then, including the three fetuin sequences reported in this work. The conserved cysteine residues are boxed and the disulphide loops they are known or predicted to form are indicated. The residues believed to be important for protease-inhibitory activity are boxed (G at position 13 and the active site QVVAG sequence, and variants thereof, see Stubbs *et al.,* 1990).

Source of sequence data:

human kininogen, Ohkubo *et al.,* (1984), Kitamura *et al.,* (1985), Kellermann *et al.,* (1986, 1987), Salvesen *et al.,* (1986); bovine kininogen, Kitamura *et al.,* (1983), Nawa *et al.,* (1983), Sueyoshi *et al.,* (1984, 1985); rat kininogen, Cole *et al.,* (1985), Furuto-Kato *et al.,* (1985), Sueyoshi *et al.,* (1985); human histidine rich glycoprotein, Koide *et al.,* (1986); human α_2-HS glycoprotein, Yoshioka *et al.,* (1986), Lee *et al.,* (1987a); rat cystatin α(A), Takio *et al.,* (1984); rat cystatin β(B), Takio *et al.,* (1983), oryzacystatin I; Abe *et al.,* (1987); human cystatin C, Grubb & Löfberg, (1982, 1985), Grubb *et al.,* (1984); bovine colostrum cysteine protease inhibitor (CPI), Hirado *et al.,* (1985); human cystatin SN, Isemura *et al.,* (1986), Al-Hashimi *et al.,* (1988), Saitoh *et al.,* (1988); chicken cystatin; Turk *et al.,* (1983), Grubb *et al.,* (1984), Schwabe *et al.,* (1984); bovine, ovine and porcine fetuins; Figs. 15, 18, 20; rat fetuin (pp63), Auberger *et al.,* (1989); mouse fetuin, Yang *et al.,* (1991).


1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7
C loop
D loop
active site
KININOGEN DOMAIN I
human, 9-131 l c s r l l l s l t Q E S Q S E E I D C N D K D L F K A V D A A L K - - K Y N S Q N Q S N N Q F V L Y R I T E A T - - K T V G S D T F
rat, 9-130 l c s r l l p s l a Q E E G A Q E L N C N D E T V F Q A V D T A L K - - K Y N A E L E S G N Q F V L Y R V T K G T - - K K D G A E T L
bovine, 9-130 l c s r l l p s l t Q E E - S Q E I D C N D Q D V F K A V D A A L T - - K Y N S E N K S G N Q F V L Y R I T E V A - - R M D N P D T F
FETUIN DOMAINS I & II
human, 13-136 l w g c h s A P H G P G L I Y R Q P N C D D P E T E E A A L V A I D - - Y I N Q N L P W G Y K H T L N Q I D E V K V W P Q Q P S G E L
bovine, 13-136 l w g c h s I P L D P V A G Y K E P A C D D P D T E Q A A L A A V D - - Y I N K H L P R G Y K H T L N Q I D S V K V W P R R P T G E V
sheep, 13-136 l c s c r s I P L D P I A G Y K E P A C D D P D T E Q A A L A A V D - - Y I N K H L P R G Y K H T L N Q I D S V K V W P R R P T G E V
pig, 13-136 l w g c r a V P H G P I L G Y R E P A C D D V E T E Q A A L A A V D - - Y I N K H L P R G Y K H T L N Q V D S V K V W P R R A A G E V
rat, 13-136 l w s c q s A P Q G A G L G F R E L A C D D P E T E H V A L I A V H - - Y L N K H L L Q G F R Q I L N Q I D K V K V W S R R P F G Q V
mouse, 13-136 l w g c q s A P Q G T G L G F R E L A C D D P E A E Q V A L L A V D - - Y L N N H L L Q G F K Q V L N Q I D K V K V W S R R P F G V V
human, 137-252 - D S A E D V R K V C Q D C P L L A P L N D T R V V H A A K A A L A - - A F N A Q N N G S N - F Q L E E I - S R A Q - L V P L P P S T
bovine, 137-253 - D S A E D V R K L C P D C P L L A P L N D S R V V H A V E V A L A - - T F N A E S N G S Y - L Q L V E I - S R A Q - F V P L P V S V
sheep, 137-253 - D S A E D V R K L C P D C P L L A P L N N S Q V V H A A E V A L A - - T F N A Q N N G S Y - F Q L V E I - S R A Q - F V P L P G S V
pig, 137-252 - D S A E D V H K V C P N C P L L A P L N D S R V V H A A E S A L A - - A F N A Q S N G S Y - L Q L V E I - S R A Q - L V P L S A S V
rat, 137-252 - D S A E D V R K F C P R C P I L I R F N D T N V V H T V K T A L A - - A F N A Q N N G T Y - F K L V E I - S R A Q - N V P F P V S T
mouse, 137-252 - D S A E D V R K L C P R C P L L T P F N D T N V V H T V N T A L A - - A F N T Q N N G T Y - F K L V E I - S R A Q - N V P L P V S T
HISTIDINE RICH GLYCOPROTEIN
human, 6-130 - - a a l l l i t l q y s c a V S - P T D C A V E P E - A E K A L D D - L I N K R R R D G Y L F Q L L R I A D A H L D - R V E N T T V
human, 131-261 - S A A L A N T K D S P V L I D F F E D T E R Y R K Q - A N K A L E K K E E N D D G A S F R V D R I E R V A - R V - - - R G G E G T G
KININOGEN DOMAINS II & III
human, 132-253 E G P V V T A Q Y D C L G C V H P I S T Q S P D L E P I L R H G I Q - - Y F N N N T Q H S S L F M L N E V K - R A Q - R Q V V A G L N
rat, 131-252 K G P K K T E E D L C V G C F Q P I P M D S S D L K P V L K H A V E - - H S N N N T K H T H L F A L R E V K - S A H - S Q V V A G M N
bovine, 131-252 E G P V V T A Q Y E C L G C V H P I S T K S P D L E P V L R Y A I Q - - Y F N N N T S H S H L F D L K E V K - R A Q - K Q V V S G W N
human, 254-375 K D F V Q P P T K I C V G C P R D I P T N S P E L E E T L T H T I T - - K L N A E N N A T F Y F K I D N V K - K A R - V Q V V A G K K
rat, 253-375 D D L F E L L P K N C R G C P R E I P V D S P E L K E A L G H S I A - - Q L N A Q H N H I F Y F K I D T V K - K A T - S Q V V A G V I
bovine, 253-372 E D F - - L P P M V C V G C P K P I P V D S P D L E E A L N H S I A - - K L N A E H D G T F Y F K I D T V K - K A T - V Q V V G G L K
CYSTATINS
chicken - - - - S E D R S R L L G A P V P V D E N D E G L Q R A L Q F A M A - - E Y N R A S N D K Y S S R V V R V I - S A K - R Q L V S G I K
rat cys C - - G T S R P P P R L L G A P Q E A D A S E E G V Q R A L D F A V S - - E Y N K G S N D A Y H S R A I Q V V - R A R - K Q L V A G I N
human, SN - - - - - - - - - - I I P G G I Y N A D L N D E W V Q R A L H F A I S - - E Y N K A T K D D Y Y R R P L R V L - R A R - Q Q T V G G V N
human, C - - S S P G K P P R L V G G P M D A S V E E E G V R R A L D F A V G - - E Y N K A S N D M Y H S R A L Q V V - R A R - K Q I V A G V N
bovine colost CPI - - - - - - - - - - R L L G G L M E A D V N E E G V Q E A L S F A V S - - E F N K R S N D A Y Q S R V V R V V - R A R - K Q V V S G M N
rat cys α - - - M D P G T T G I V G G V S E A K P A T P E I Q E V A D K V K R - - Q L E E K T N E K Y - - E K F K V V - E Y K - S Q V V A G Q I
rat cys β - - - - - - - - - - M M C G A P S A T M P A T T E T Q E I A D K V K S - - Q L E E K A N Q K F - - D V F K A I - S F R - R Q V V A G T N
oryzacystatin I M S S D G G P V - - L G G V E P V G N E N D L H L V D L A R F A V T - - E H N K K A N S L L - - E F E K L V S V K - Q Q V V A G T L


Cystatin A loop
Cystatin B loop
KININOGEN DOMAIN I
human, 9-131 Y S F K Y E I K E G D C P V Q S G K T W - Q D C E Y K D A A K A A - - - T G E C T A - T V G K R S S T K - - F S V A T Q - Y C Q I T P A
rat, 9-131 Y S F K Y Q I K E G N C S V Q S G L T W - Q D C D F K D A E E A A - - - T G E C T T - T L G K K - E N K - - F S V A T Q - I C N I T P G
bovine, 9-130 Y S L K Y Q I K E G D C P F Q S N K T W - Q D C D Y K D S A Q A A - - - T G Q C T A - T V A K R G N M K - - F S V A I Q - T C L I T P A
FETUIN DOMAINS I & II
human, 13-136 F E I E I D T L E T T C H V L D P T P V - A R C S V R Q L K E H A - - V E G D C D F - - Q L L K L D G K - - F S V V Y - A K C D - S S P
bovine, 13-136 Y D I E I D T L E T T C H V L D P T P L - A N C S V R Q Q T Q H A - - V E G D C D I - - H V L K Q D G Q - - F S V L F - T K C D - S S P
sheep, 13-136 Y D I E I D T L E T T C H V L D P T P L - V N C S V R Q Q T E H A - - V E G D C D I - - H V L K Q D G Q - - F S V L F - T K C D - S S P
pig, 13-136 F D I E I D T L E T T C H V L D P T P L - A N C S V R Q L T E H A - - V E G D C D F - - H V L K Q D G Q - - F S V L F - T K C D - S S P
rat, 13-136 Y E L E I D T L E T T C H A L D P T P L - A N C S V R Q Q A E H A - - V E G D C D F - - H I L K Q D G Q - - F R V L H - A Q C H - S T P
mouse, 13-136 Y E M E V D T L E T T C H A L D P T P L - A N C S V R Q L T E H A - - V E G D C D F - - H I L K Q D G Q - - F R V M H - T Q C H - S T P
human, 137-252 Y - V E F T V S G T D C V A K E A T E A - A K C - - N L L A E K Q - - - Y G F C K A - T L S E K - L G G - - A E V A - - V T C T V F Q T
bovine, 137-253 S - V E F A V A A T D C I A K E V V D P - T K C - - N L L A E K Q - - - Y G F C K G - S V I Q K A L G G - - E D V R - - V T C T L F Q T
sheep, 137-253 S - V E F A V A A T D C I A K E V V D P - T K C - - N L L A E K Q - - - Y G F C K G - S V I Q K A L G G - - E D V T - - V T C T L F Q T
pig, 137-252 S - V E F A D A V T D C V A K E A Y S P - T K C - - N L L V E K Q - - - Y G F C K G - T V T A K - V N E - - E D V A - - V T C T V F Q T
rat, 137-252 L - V E F V I A A T D C T G Q E V T D P - A K C - - N L L A E K Q - - - Y G F C K A - T L I H R - L G G - - E E V S - - V A C K L F Q T
mouse, 137-252 L - V E F V I A A T D C T A K E V T D P - A K C - - N L L A E K Q - - - H G F C K A - N L M H N - L G G - - E E V S - - V A C K L F Q T
HISTIDINE RICH GLYCOPROTEIN
human, 6-130 Y Y L V L D V Q E S D C S V L R K Y W N D - - C E P P D S R R P S E I V I G Q C K V I A T R H S H E S Q D L - - R V I D F N C T T S S V
human, 131-261 Y F V D F S V R N - - - - - - - - - - - - - C P R H H F P R H P N - V F G F C R A D L F Y D V E A L D L E S P K N L V I N C E V F D -
KININOGEN DOMAINS II & III
human, 132-253 F R I T Y S I V Q T N C S K E N F L F L T P D C - - K S L W N G D - - - T G E C T D N A Y I D I Q L R I A S F S - - - - Q N C D I Y P G
rat, 131-252 Y K I I Y S I V Q T N C S K E D F P S L R E D C - - V P L P Y G D - - - H G E C T G H T H V D I H N T I A G F S - - - - Q S C D L Y P G
bovine, 131-252 Y E V N Y S I A Q T N C S K E E F S F L T P D C - - K S L S S G D - - - T G E C T D K A H V D V K L R I S S F S - - - - Q K C D L Y P G
human, 254-375 Y F I D F V A R E T T C S K E S N E E L T E S C - - E T K K L G Q - - - S L D C N A E V Y Y V V P W E K K I Y P T - - - - V N C Q P L G M
rat, 253-375 Y V I E F I A R E T N C S K Q S K T E L T A D C - - E T K H L G Q - - - S L N C N A N V Y M R P W E N K V V P T - - - - V R C Q A L D M
bovine, 253-372 Y S I V F I A R E T T C S K G S N E E L T K S C - - E I N I H G Q - - - I L H C D A N V Y Y V V P W E E K V Y P T - - - - V N C Q P L G Q
CYSTATINS
chicken Y I L Q V E I G R T T C P K - S S G D L - Q S C E F H D E P E M A - - K Y T T C T F V V Y S I P W L N Q I K L L - - - E S K C Q - - - -
rat cys C Y Y L D V E M G R T T C T K - S Q T N L - T N C P F H D Q P H L M - - R K A L C S F Q I Y S V P W K G T H T L T - - - K S S C K N A - -
human, SN Y F F D V E V G R T I C T K - S Q P N L - D T C A F H E Q P E L Q - - K K Q L C S F E I Y E V P W E N R R S L V - - - K S R C Q E S - -
human, C Y F L D V E L G R T T C T K - T Q P N L - D N C P F H D Q P H L K - - R K A F C S F Q I Y A V P S Q G T M T L S - - - K S T C Q D A - -
bovine colost CPI Y F L D V E L G R T T C T K - S Q A N L - D S C P F H N Q P H L K - - R E K L C S F Q V Y Y V V P W M N T I N L V - - - K F S C Q D - - -
rat cys α M K V D V G N G R F L H M K - - - - - - - - - - V L R G L S G D D - - - - D L - K L L D Y Q T N K T K N D E L T D F - - - - - - - - - - -
rat cys β I K V D V G E E - - K C V H L R - - - - - - - - V F E P L P H E N - - K - P L - T L S S Y Q T D K E K H D E L T Y F - - - - - - - - - - -
oryzacystatin I Y Y F T I E V K E G D - A K K L Y E A K - V W E K P W M D F K E L Q E F K P V D A S A N A -

The disulphide loop structure predicted for human α_2-HS glycoprotein, on the basis of its homology to the cystatins, and, more importantly, to the cystatin domains in the kininogens (Elzanowski *et al.,* 1988) was soon confirmed by protein chemistry (Araki *et al.,* 1989; Kellermann *et al.,* 1989). Both papers reported the presence of linearly arranged and tandemly repeated disulphide loops within the larger of the α_2-HS glycoprotein chains, and a single interchain disulphide bond joining the 2 chains (see Fig. 7).

All of the known mammalian fetuin sequences also contain 12 cysteine residues in identical positions to those in human α_2-HS glycoprotein (Fig. 24). While the disulphide loop structure has not been established for any other fetuin, the high degree of sequence identity, overall, and in the regions adjacent to the cysteines, strongly suggests that the disulphide structure identified in human α_2-HS glycoprotein is a common feature of the fetuin family. In support of this, Marti *et al.,* (1973) and Spiro, (1963) noted that sheep and bovine fetuin, respectively, contained 12 cysteine residues, but no free -SH groups, implying six disulphide bonds.

4.5. Gene organisation within the cystatin superfamily.

4.5.1. Organisation of the fetuin gene.

The proposed three domain structure of the fetuin protein is precisely reflected in the organisation of the rat fetuin (pp63) gene which has recently been determined (Falquerho *et al.,* 1991). The two cystatin-like domains, D1 and D2, are each encoded by three exons separated by two introns, as are cystatin domains in other members of the superfamily: cystatins (see for example Saitoh *et al.,* 1989; Abrahamson *et al.,* 1990), HRG (Koide, 1988) and kininogens (see for example Nakanishi *et al.,* 1987; Anderson *et al.,* 1989). The stefins, too, are encoded by three exons, but in a markedly different arrangement, see Fig. 28. The terminal domain, D3, is encoded by a single exon, as is the non-cystatin carboxyl-terminal domain in the kininogens. However in high and low

molecular weight kininogens two alternative carboxyl-terminal domains are encoded by alternative last exons. Furthermore, as pointed out by Elzanowski *et al.,* (1988), the carboxyl-terminal sequence immediately surrounding the last cysteine residue *(pro-pro-CYS-pro-gly-arg* in α_2-HS glycoprotein, see Fig. 24) shows significant homology with the same region of human *(pro-lys-CYS-pro-gly-arg)* and bovine *(pro-lys-CYS-pro-ser-arg)* high molecular weight kininogen, suggesting a possible common origin of the carboxyl-terminal regions of the two proteins.

Fig. 27. The organisation of the rat fetuin (pp63) gene.

The domain structure proposed on the basis of the disulphide loop arrangement and sequence comparisons is exactly reflected in the organisation of the rat fetuin gene. The signal peptide and cystatin domain D1 are encoded by exons I to III, the second cystatin domain is encoded by exons IV to VI and the terminal domain D3 is encoded entirely by exon VII. The bottom scale is in kb from the transcriptional start point identified by Falquerho *et al.,* (1991). Lengths of introns and exons are approximately to scale. Open boxes denote coding regions, solid lines denote introns. The protein is represented as an open bar, as in Fig. 25. The disulphide loop structure is indicated above this.

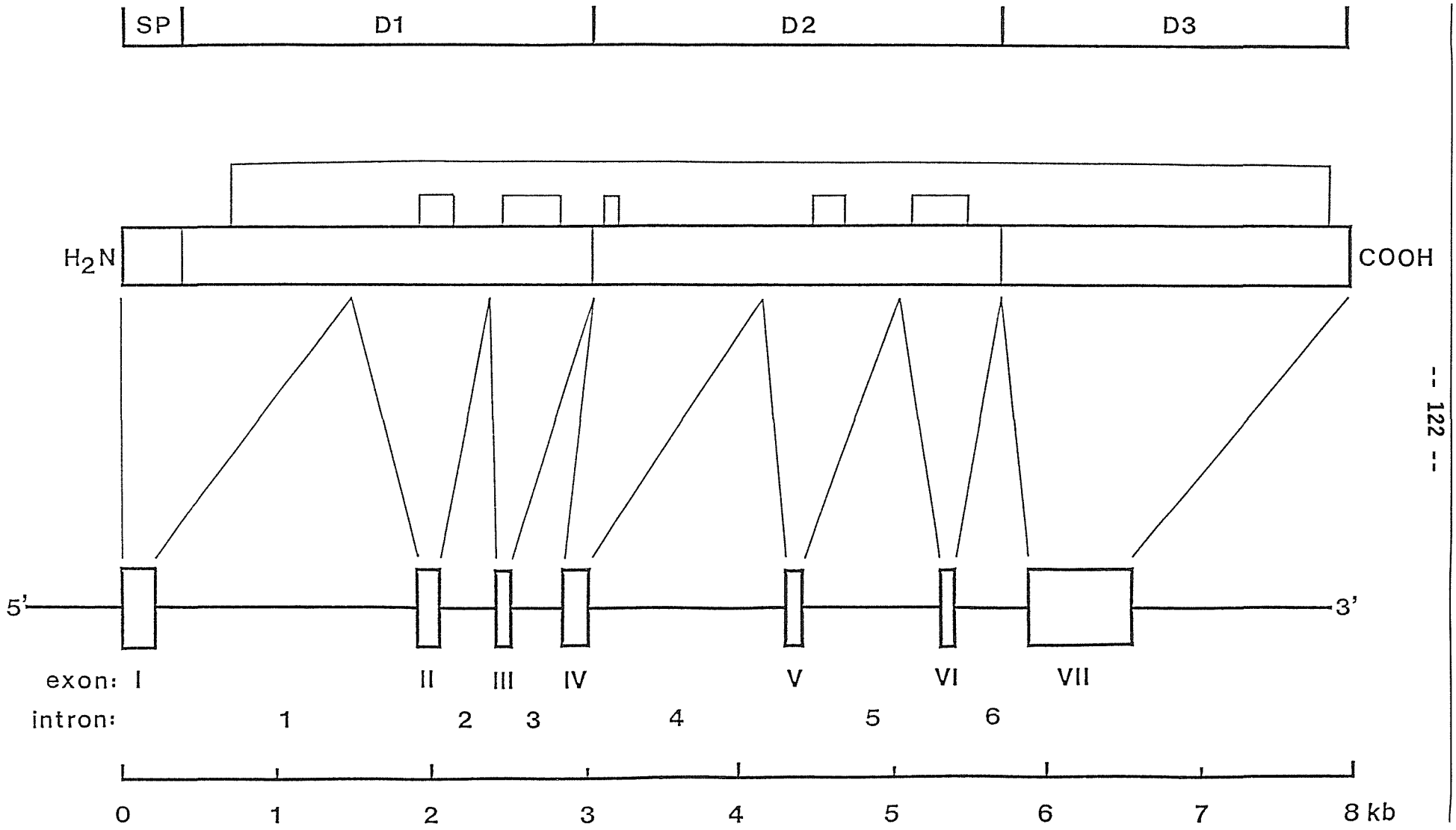
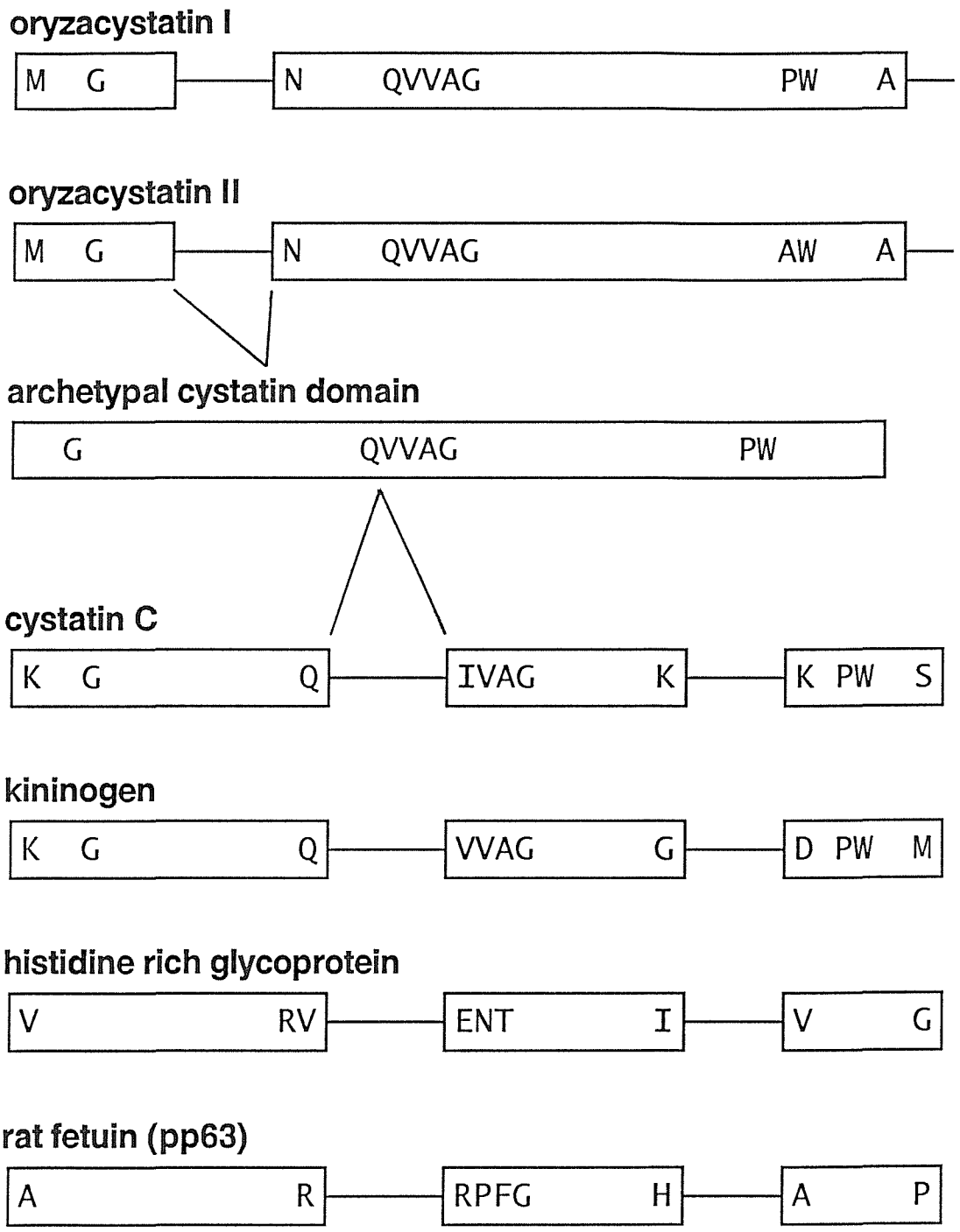

SP
D1
D2
D3
H2N
COOH
5'
3'
exon: I II III IV V VI VII
intron: 1 2 3 4 5 6
0 1 2 3 4 5 6 7 8 kb

Fig. 28. Gene organisation in the cystatin superfamily.

The gene organisation of the cystatin domains in the various members of the cystatin superfamily is displayed diagrammatically. Lengths of introns and exons are *not* to scale. Open boxes denote coding regions, solid lines denote introns, amino acids indicated are in single letter code. Residues strongly conserved throughout the superfamily (G, *gly,* at or near position 10, the active site QVVAG sequence and variants thereof, and the PW, *pro-trp,* sequence, see also Fig. 26) are indicated, as are the first and last amino acids in each cystatin domain. As discussed in the text, in every member of the superfamily for which the genomic structure is known, the cystatin domain comprises 3 exons interrupted by two introns. In the case of the oryzacystatins, the third intron interrupts the 3'-UTR of the cDNA. This figure is modified from Fig. 3, Kondo *et al.,* (1991).

Source of data: oryzacystatin I, Kondo *et al.,* (1989); oryzacystatin II, Kondo *et al.,* (1991); cystatin C (human), Saitoh *et al.,* (1989); kininogen (human, domain D3), Ohkubo *et al.,* (1984), Kitamura *et al.,* (1985), Kellermann *et al.,* (1986, 1987), Salvesen *et al.,* (1986); histidine rich glycoprotein (human, domain D1), Koide, (1988), rat fetuin (pp63, domain D1), Falquerho *et al.,* (1991).

Fig. 28. Gene organisation in the cystatin superfamily.



4.6. Evolution of the cystatin superfamily.

As discussed above, the cystatins are low molecular weight cysteine protease inhibitors with homology to chicken cystatin, and the cystatin *superfamily* was born when it was first seen that multiple copies of a cystatin-like sequence were present in the kininogens. Furthermore, it is now known that cystatin-like sequences are present in the fetuins and histidine-rich glycoprotein. Over recent years a series of papers have discussed the evolution of the cystatin superfamily, the most recent and comprehensive treatise on the subject being that of Rawlings & Barrett, (1990); see also Müller-Esterl *et al.,* (1985); Barrett *et al.,* (1986); Turk *et al.,* (1986); Elzanowski *et al.,* (1988).

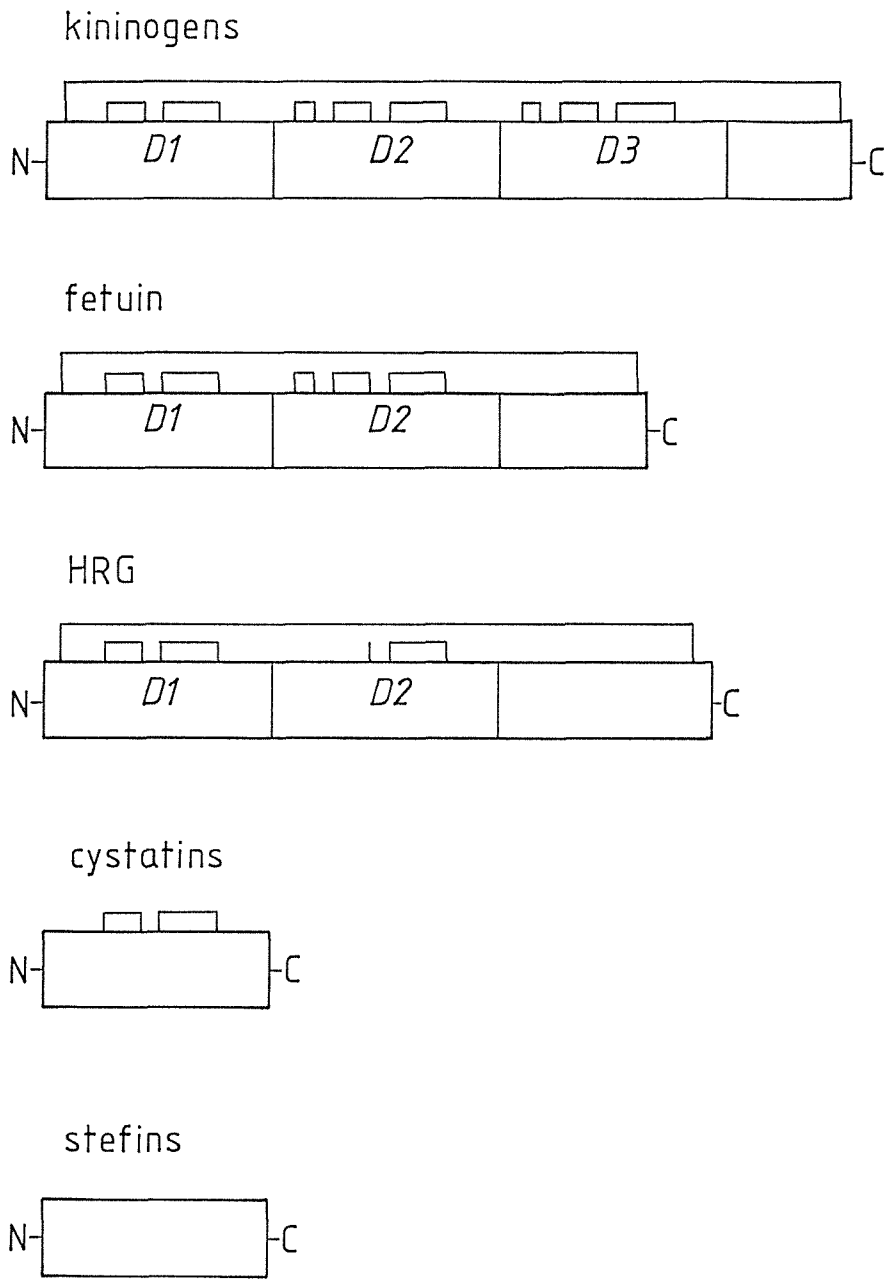
Müller-Esterl *et al.,* (1985) proposed a scheme for the evolution of the various members of the cystatin superfamily, in which the cystatins, and from them the kininogens, arose from a precursor stefin-like protein by a fusion of separate exons, of which only the one encoding the N-terminal sequence was related to the stefins. This hypothesis can, however, be ruled out on two counts:
1) analysis of sequence alignments reveals that sequences are common between the stefins and the cystatins throughout the length of the molecules and *not* just at the amino terminus; for example the PW *(pro-trp)* sequence (residues 115-116, Fig. 26) and a strongly conserved tyrosine (residue 112, Fig. 26) found towards the carboxyl-terminus of the cystatin domain are common to oryza-cystatin I (a stefin), many cystatins and kininogen domain D3 (see Figs. 26, 28 and Rawlings & Barrett, 1990).
2) the recently-determined crystal structure of human stefin B and comparison with the known crystal structure of chicken cystatin (see Stubbs *et al.,* 1990) supports very much what was said in 1). Indeed, on the basis of the two crystal structures, Stubbs *et al.,* (1990) proposed a sequence alignment on the basis of the topographical equivalence of the residues in the two structures.

Fig. 29. The stefin/cystatin sequence alignment proposed by Stubbs *et al.*, (1990) on the basis of the topographical equivalence of the residues. Human stefin B is aligned with chicken cystatin. The active site QVVAG and the PW sequence mentioned above are boxed. The disulphide loops in the cystatin sequence are indicated.

```
stefin B   - - - - - M M C G A P S A T Q P A T A E T Q H I A D Q V R S Q L E E K Y N K K F - -
cystatin   S E D R S R L L G A P V P V D E N D E G L Q R A L Q F A M A E Y N R A S N D K Y S S

stefin B   P V F K A V S F K S [Q V V A G] T N Y F I K V H V G D - - - - - - - - - - - - - - -
cystatin   R V V R V I S A K R [Q L V S G] I K Y I L Q V E I G R T T [C] P K S S G D L Q S [C] E F H

stefin B   - - - - - - - E D F V H L R V F Q S L [P H] E N K P L T L S N Y Q T N K A K H D E L T
cystatin   D E P E M A K Y T T [C] T F V V Y S I - [P W] L - N Q I K L L E S K [C] - Q - - - - - -

stefin B   Y F *
cystatin   - - *
```

Fig. 30. Gross structure of cystatin superfamily members.
The gross structure of the proteins of the cystatin superfamily is shown. The cystatin domains are numbered from the amino-terminus. The disulphide loop structure is indicated above the open bar which represents the polypeptide chain. The disulphide loop structure for HRG is inferred from sequence comparisons. It has not been experimentally determined.

Fig. 30. Gross structure of the cystatin superfamlly members.



A further spanner has been thrown into the works of the evolutionary story, by the discovery of Kondo *et al.*, (1989, 1991, see Fig. 28), that despite the high protein sequence homology, the genomic organisation of the oryzacystatins is markedly different to that seen in the animal cystatins. Indeed, Kondo *et al.*, (1991) go on to suggest that the oryzacystatins belong to a new phytocystatin family within the cystatin superfamily.

Clearly, more work is needed to complete this evolutionary pathway, in particular animal stefin genomic sequences are required. From what is already known, it is clear that the rice oryzacystatins and the animal cystatins diverged from a common ancestor before the animal cystatins themselves diverged and gave rise to the larger proteins.

4.7. Analysis of the rat fetuin gene promoter region.

The 5'-regulatory region of the rat fetuin (pp63) gene has recently been cloned and sequenced (Falquerho *et al.*, 1991). Analysis of the sequence reported (Genbank accession number M36547) reveals a series of potential *cis*-elements, *i.e.* binding sites for sequence-specific DNA-binding proteins or transcription factors. Falqherho *et al.*, (1991) identified a (putative) TATA box and C/EBP, octamer and NF-1 binding sites. These are indicated on Fig. 35. In addition to these, several potential IL-6 responsive elements and putative binding sites for a number of other factors are readily apparent.

4.7.1. Interleukin-6-responsive elements.

Analysis of the promoter regions of acute phase genes has enabled the identification of several types of IL-6-responsive elements (see Fig. 3). Furthermore deletion and reporter gene analysis, and, more recently, the cloning and sequencing of several transcription factors that actually bind these sites has confirmed their involvement, if not their rôle, in the changes in the rate of transcription observed in the acute phase response. Recently it has been reported that rat fetuin is a negative acute phase protein, and control of expression was at the level of transcription (Daveau *et al.*, 1990). In the light of recent results (Haasemann *et al.*, 1991; Brown *et al.*, 1991a) it now seems certain that the mRNA Daveau *et al.*, (1990) were measuring (with a human α_2-HS glycoprotein cDNA probe) was that of pp63. Examination of the rat fetuin promoter sequence reveals a series of potential type A and B IL-6-responsive elements resembling the consensus sequences, see Fig. 31.

Fig. 31. IL-6-responsive elements in the rat fetuin gene promoter.

Gene	Location	Sequence Observed	Reference
Type A elements			
rat fetuin	-212	A T G T T G C A A G	Falquerho *et al.*, (1991)
	-166	C T T T G G C A A A	
	-157	T G T T T G C A A G	
	-61	T T T A C G C A A T	
	CONSENSUS	TTNNGNAAT G	
Type B elements			
rat fetuin	-791	T T C A G G G A G	Falqherho *et al.*, (1991)
	-728	T T C A G G G A G	
	-667	C T T T A G G G A	
	-646	T T C A G G G A A	
	-428	T T C T A G G A G	
	-164	C T T T G G C A A	
	CONSENSUS	TTCTGGGAA T A T	

4.7.2. TPA-responsive element.

TPA (12-O-tetradecanoyl-phorbol-13-acetate) and other phorbol esters are potent tumour promoters (see Weinstein *et al.*, 1979; Slaga, 1983). Transcription from a number of genes has been demonstrated to be TPA-inducible: proto-oncogenes c-*fos* (Greenberg & Ziff, 1984; Kruijer *et al.*, 1984), c-*myc* (Kelly *et al.*, 1983; Greenberg & Ziff, 1984), c-*sis* (Colamonici *et al.*, 1986) and pro-1 (Lerman & Colburn, 1987), and a series of other genes including collagenase (Whitham *et al.*, 1986) and stromelysin (Matrisian *et al.*, 1986; Whitham *et al.*, 1986). Recent analysis of the human collagenase gene (Angel *et al.*, 1987) has revealed a small *cis*-element which confers TPA-inducibility upon both the human collagenase gene and reporter genes. Footprinting experiments (Angel *et al.*, 1987) further demonstrated that the *cis*-element identified was the binding site for the transcription factor AP-1 (Lee *et al.*, 1987b, 1987c). Examination of the promoter regions of TPA-inducible genes has shown that this sequence is well conserved (see Fig. 32) and a very similar sequence can be seen in the rat fetuin promoter at positions -472 to -463 (see Fig. 32).

Fig. 32. TPA-responsive element sequences..

Gene	Location	Sequence Observed	Reference
human collagenase	-72	A T G A G T C A G A	Angel *et al.*, (1987)
rat stromelysin	-71	A T G A G T C A G T	Matrisian *et al.*, (1986)
human MTIIA	-104	G T G A C T C A G T	Karin & Richards, (1982)
SV40	120	A T T A G T C A G C	Buchman *et al.*, (1980)
	30	A T T A G T C A G C	
polyoma virus	5129	C T G C G T C A G T	Griffin *et al.*, (1980)
	5121	G T T A G T C A C T	Zenke *et al.*, (1986)
interleukin 2	-185	T T C A G T C A G T	Fujita *et al.*, (1983)
rat fetuin	-472	A T G A C T C A C T	Falquerho *et al.*, (1991)
	CONSENSUS	TGAGTCAG	

4.7.3. Heat-shock element.

When subjected to increased temperatures, cells rapidly synthesise large amounts of a limited number of proteins (the HSP's or heat-shock proteins) while reducing the synthesis of other proteins (for review see Lindquist, 1986). The heat shock response is believed to be evolutionarily universal, and has been most heavily studied in *Drosophila*, see Ashburner & Bonner, (1979). In *Drosophila*, there are seven major heat-shock proteins and the promoter regions of the genes were all cloned and sequenced in the early 1980's (Karch *et al.*, 1980; Török & Karch, 1980; Ingolia *et al.*, 1980; Ingolia & Craig, 1981; Holmgren *et al.*, 1981). Analysis of the 5'-regulatory regions of these genes revealed a rather poorly conserved 14bp element, upstream of the TATA box. That the sequence identified was genuinely involved was again demonstrated by ligating it to a heterologous gene and showing that the construct was heat-shock inducible (Pelham, 1982). Beyond this conserved sequence, there is considerable variation in the promoter organisation of heat-shock genes; all the sequence information required for hsp70 induction are present within 90bp of the transcription start point (tsp), whereas over 500bp 5' to the tsp are required for hsp23 (see Thomas & Elgin, 1988 for references and discussion). A sequence resembling the consensus heat-shock *cis*-element can be seen in the rat fetuin promoter, at position -520 to -506, see Figs. 33 and 35. Note, however that is considerably further from the transcriptional startpoint than are the example sequences in Fig. 33.

Fig. 33. Heat-shock element sequences.

Gene	Location	Sequence Observed	Reference
hsp 70	-28	C T C G A A T G T T C G C G	Ingolia *et al.*, (1980)
hsp 83	-41	C T A G A A G T T T C T A G	Holmgren *et al.*, (1981)
hsp 22	-326	C C G G T A T T T T C T A G	Ingolia & Craig, (1981)
hsp 26	-29	C C G G A C T C T T C T A G	Holmgren *et al.*, (1981)
hsp 27	-188	T T G C C A T G C A C T A G	Ingolia & Craig, (1981)
hsp 68	-37	C A G G G A A A T C T C G A	Holmgren *et al.*, (1981)
hsp 68	-28	C T C G A A T T T T C C C C	Holmgren *et al.*, (1986)
hsp 23	-168	A C G C A C A C T A C G A T	Ingolia & Craig, (1981)
rat fetuin	-472	T A A G A A C T T C C T T T	Falquerho *et al.*, (1991)
	CONSENSUS	C n n G A A n n T T C n n G	

Note that numbering schemes in the papers cited differ.

4.7.4. Serum-responsive element.

It has been reported that transcription from a series of genes rises rapidly and transiently on stimulation with various growth factors purified from serum or with whole serum. Many of the genes involved remain unknown and uncharacterised (see Kelly *et al.*, 1983; Elder *et al.*, 1984; Greenberg & Ziff, 1984; Müller *et al.*, 1984; Kruijer *et al.*, 1984; Cochran *et al.*, 1984; Lau & Nathans, 1985; Imbra & Karin, 1987). Two that have been studied are the proto-oncogene c-*fos* (see Treisman, 1985, 1986; Gilman *et al.*, 1986; Prywes & Roeder, 1986; Subramaniam *et al.*, 1989; Herrera *et al.*, 1989) and several cytoskeletal actins (see Mohun *et al.*, 1987). In a manner very similar to that used to characterise the IL-6, heat-shock and TPA-responsive elements *(vide supra)*, an artificial serum-responsive element was shown to confer serum-inducibility to previously uninducible genes (Treisman, 1986). An element resembling this consensus sequence can be seen in the rat fetuin promoter, at position -330, see Figs. 34 and 35. More recently a transcription factor that binds this sequence has been characterised (Treisman, 1987).

Fig. 34. Serum-responsive elements.

Gene	Sequence Observed	Reference
c-fos^{h}	A C A G G A T G T C C A T A T T A G G A C A T C T	Treisman, (1985)
Y-actin	G A A A G A T G C C C A T A T T T G G C G A T C T	Mohun *et al.*, (1987)
rat fetuin	A T G G C A T G A T C A T A T T C A G G T A G A G	Falquerho *et al.*, (1991)
artificial	A T T G G A T G T C C A T A T T A G G A C A T C T	Treisman, (1986)

Abbreviations: c-*fos*^{h}, human c-*fos*; Y-actin, *Xenopus laevis* Y-actin; artificial, the synthetic SRE used by Treisman, (1986).

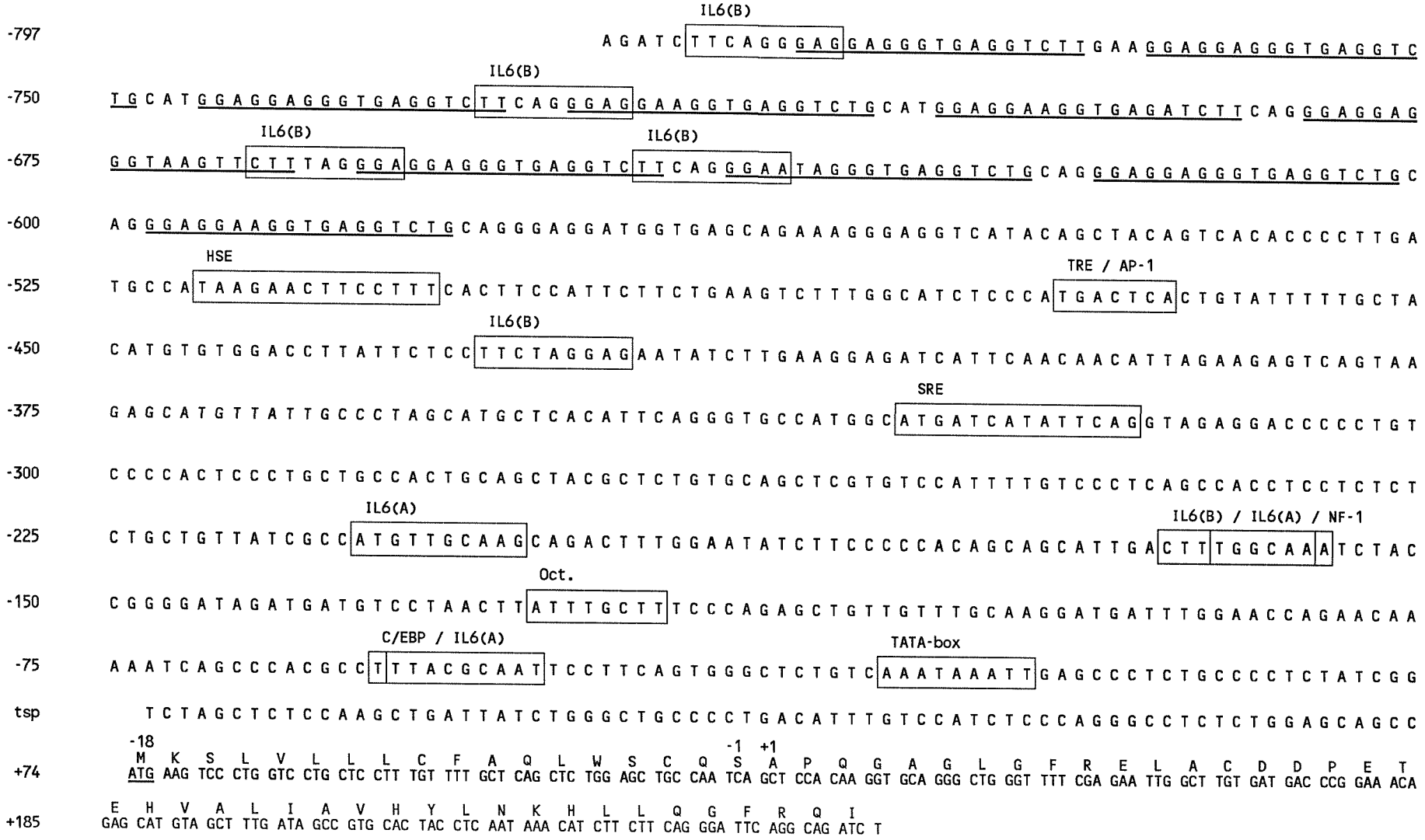
Fig. 35. The 5'-regulatory region of the rat fetuin *(pp63)* gene.

Analysis of the recently cloned and sequenced 5'-regulatory region of the rat fetuin (pp63) gene (Genbank accession number M36547) reveals a series of potential *cis*-elements, *i.e.* binding sites for sequence-specific DNA-binding proteins or transcription factors. Falquerho *et al.*, (1991) identified a (putative) TATA box and C/EBP, octamer and NF-1 binding sites, and these are indicated. In addition, the potential IL-6-, serum-, heat-shock- and TPA-responsive elements discussed in the text are indicated.

The 10-fold repeat of the sequence 5'-GGAGGAGG$^{G}_{T}$TGAGGTCT$^{T}_{G}$ with a 3 base spacer, pointed out by Falquerho *et al.*, (1991) is underlined.

tsp = transcriptional start point.

Fig. 35. The 5'-regulatory region of the rat fetuin (pp63) gene.



4.8. Fetuins and the acute phase response.

As mentioned in the Introduction, the "*acute phase response*" is the name given to the sequence of events in higher animals that follows a disturbance of homeostasis, see Fig. 1. Human α_2-HS glycoprotein is generally considered to be a negative acute phase protein (van Oss *et al.*, 1975; Lebreton *et al.*, 1979; Baskies *et al.*, 1980). More recently, in a human hepatoma cell line, it has been shown that the cytokines IL-6 and interleukin-1β caused the down-regulation of α_2-HS glycoprotein synthesis and that this decrease could be observed at the protein and mRNA level (Daveau *et al.*, 1988). It has also been demonstrated that rat fetuin is a negative acute phase protein, and that control of its expression is at the level of transcription (Daveau *et al.*, 1990). Potential IL-6 responsive elements can be identified in the rat fetuin (pp63) gene regulatory region (see Fig. 31).

As discussed in the Introduction, several types of IL-6 responsive elements have been identified in the 5'-regulatory regions of a series of acute phase genes (see Won & Baumann, 1990, and Fig. 3). More recently several sequence-specific DNA binding proteins have been characterised and cloned that have been shown to bind to these IL-6 responsive elements (Akira *et al.*, 1990; Poli *et al.*, 1990).

Recently we have shown that in cattle, fetuin is a positive acute phase protein, its level in plasma increasing by up to an order of magnitude in some trauma cases (see Dziegielewska *et al.*, 1991). Similar species variability has been seen for transferrin, it being a positive acute phase protein in rodents and a negative one in man. Clearly in both cases (transferrin and fetuin) it is very difficult to envisage the function of these (opposite) changes.

Figs. 36 and 37. Bovine fetuin, a new acute phase protein.

Cattle serum samples were divided into four groups; normal, pregnant, post-abortion and other trauma. Concentrations of albumin and fetuin were measured by the radial immunodiffusion method of Mancini *et al.,* (1965). Total protein was measured using the Bradford method (Bradford, 1976). The two figures, 36 and 37, are redrawn from Dziegielewska *et al.,* (1991).

Fig. 36. Fetuin as a proportion of total plasma protein.
The serum concentration of fetuin, as a proportion of total protein is expressed as mean values ∓ standard errors of the mean (error bars).

Fig. 37. Scatterplot of serum fetuin vs serum albumin concentrations.
The serum concentration of fetuin is plotted against that of albumin for each animal. The simple linear regression line for all data is shown. There is a highly significant negative correlation between fetuin and albumin concentrations, implying that in cattle fetuin is a positive acute phase protein.

Fig. 36. Fetuin as a proportion of total plasma protein.

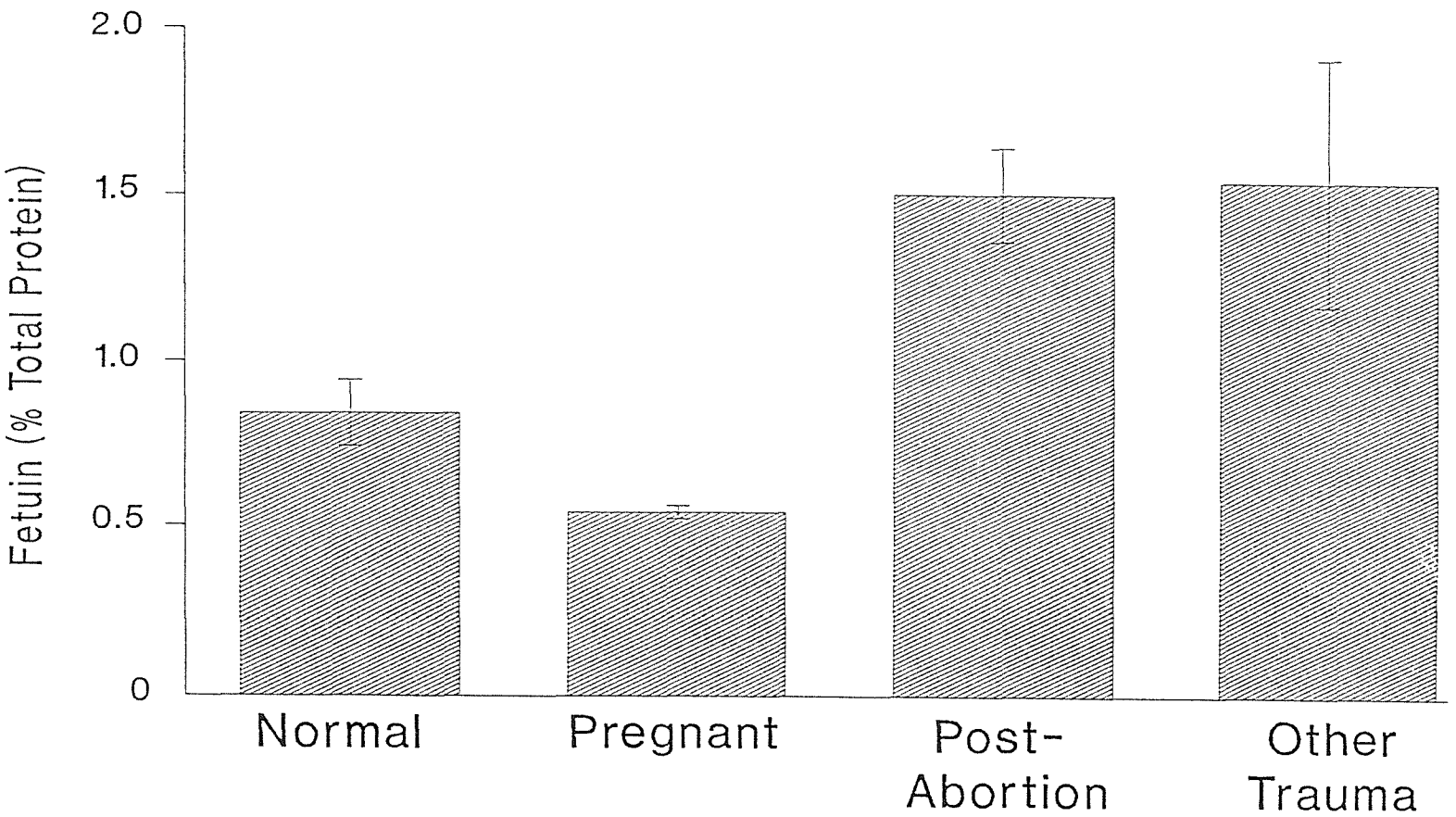
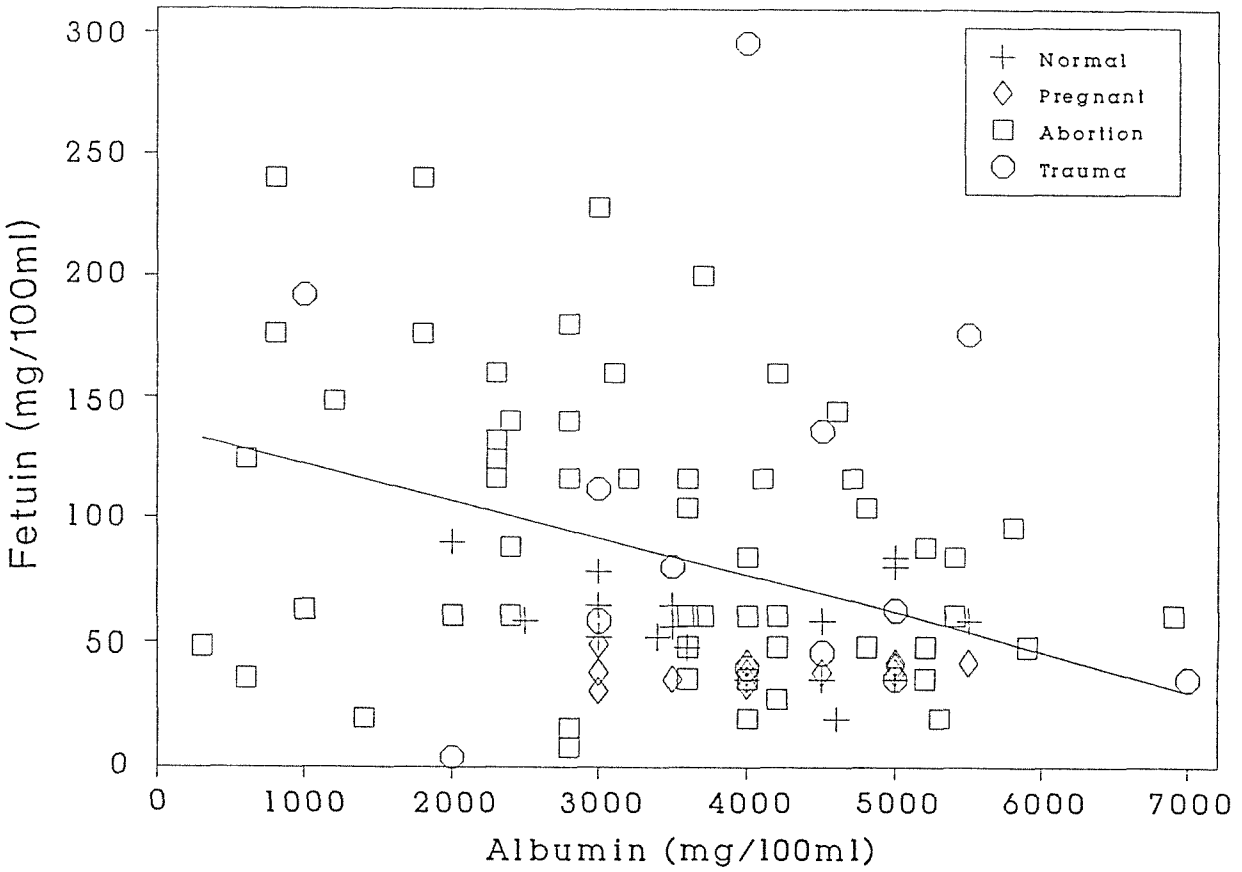


Fig. 37. Scatterplot of serum fetuin vs serum albumin concentrations.



4.9. Fetuin function.

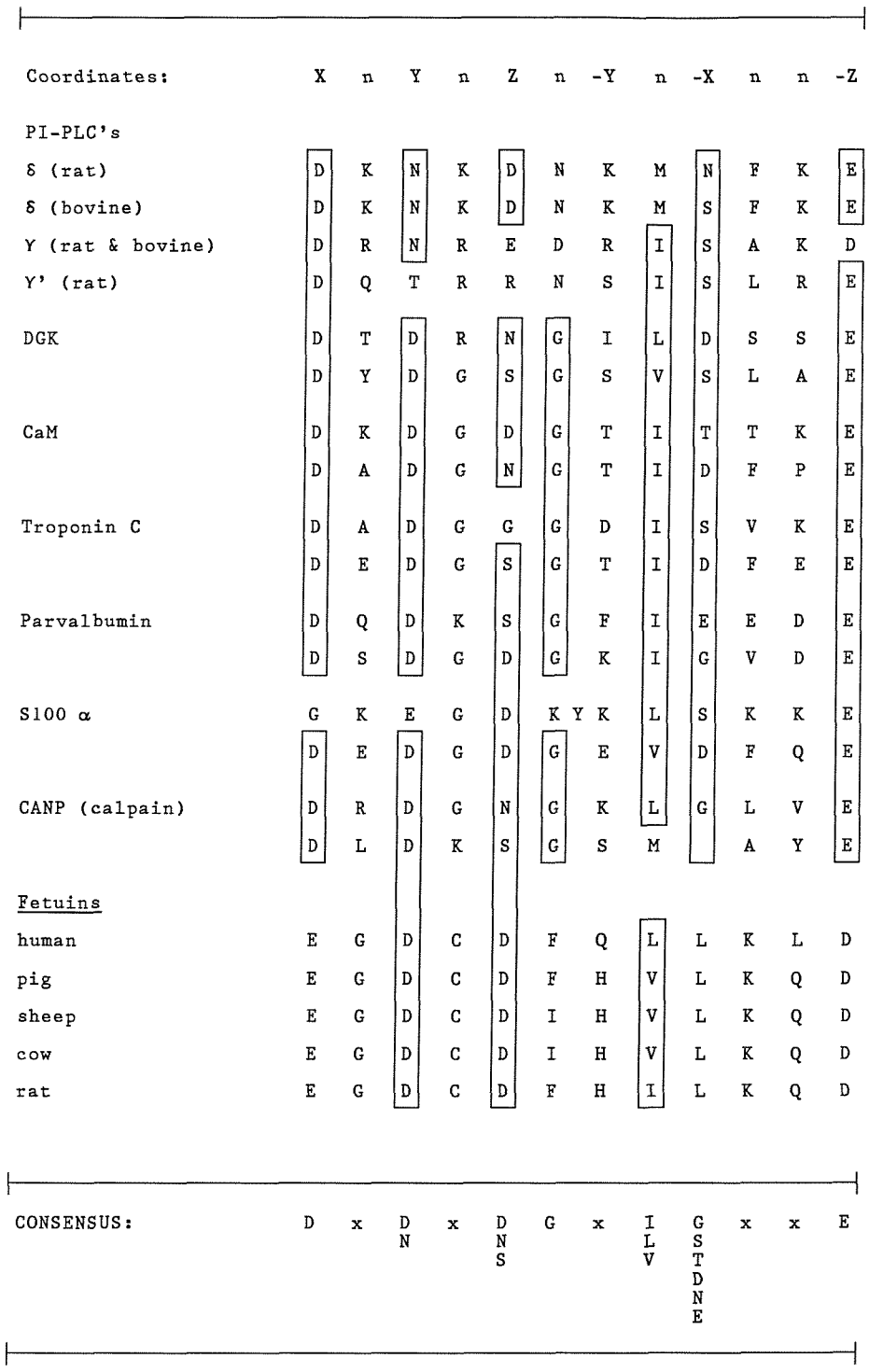
While the cystatin superfamily clearly comprises a group of related proteins containing tandem repeats of cystatin-like domains, it appears most unlikely that the function of all of these proteins is similar, and the precise functional rôle of these structurally defined motifs remains obscure. Studies dealing with purified fetuins, in collaboration with Prof. W. Müller-Esterl, have shown that they are at best very weak inhibitors of papain or cathepsins, thus removing the possibility that the amino-terminal cystatin-like domains (D1 and D2) are functional cysteine protease inhibitors. This is perhaps not surprising, bearing in mind the sequence alignment shown in Fig. 26, and the residues now known to be involved in the interaction with the cysteine protease (see Stubbs *et al.*, 1990).

A potential function that could be attributed to the cystatin domain(s) is the Ca^{2+}-binding property of fetuins, demonstrated for the human protein (Wilson *et al.*, 1977). The cystatin-like domain D1 of human kininogens - which has lost its protease inhibitory activity - does bind Ca^{2+} ions weakly (Ishiguro *et al.*, 1988). Fetuin domain D1 contains a sequence motif resembling the consensus sequence of EF hand motifs (see Kretsinger, 1987; Bairoch & Cox, 1990)

Fig. 38. Example EF-hand motifs from the literature aligned with the sequences identified in the D1 domain of the fetuins.
Abbreviations and source of data: PI-PLC, phospholipid specific phospholipase C, Bairoch & Cox, (1990); DGK, porcine diacylglycerol kinase, Sakane *et al.*, (1990); CaM, chicken calmodulin, Putkey *et al.*, (1983); Troponin C (rabbit), Wilkinson *et al.*, (1980); parvalbumin (carp), Tufty & Kretsinger, (1975); S100 α, bovine S100 protein, α-subunit, Kuwano *et al.*, (1986); CANP, rabbit calcium protease (μ type, CANP or calpain I), Emori *et al.*, (1986).

X,Y,Z,-Y,-X,-Z mark the Ca^{2+} ion coordinating residues. "x" = any amino acid.

Fig. 38. EF hand motifs.



Coordinates:	X	n	Y	n	Z	n	-Y	n	-X	n	n	-Z
PI-PLC's												
δ (rat)	D	K	N	K	D	N	K	M	N	F	K	E
δ (bovine)	D	K	N	K	D	N	K	M	S	F	K	E
Y (rat & bovine)	D	R	N	R	E	D	R	I	S	A	K	D
Y' (rat)	D	Q	T	R	R	N	S	I	S	L	R	E
DGK	D	T	D	R	N	G	I	L	D	S	S	E
	D	Y	D	G	S	G	S	V	S	L	A	E
CaM	D	K	D	G	D	G	T	I	T	T	K	E
	D	A	D	G	N	G	T	I	D	F	P	E
Troponin C	D	A	D	G	G	G	D	I	S	V	K	E
	D	E	D	G	S	G	T	I	D	F	E	E
Parvalbumin	D	Q	D	K	S	G	F	I	E	E	D	E
	D	S	D	G	D	G	K	I	G	V	D	E
S100 α	G	K	E	G	D	K	Y K	L	S	K	K	E
	D	E	D	G	D	G	E	V	D	F	Q	E
CANP (calpain)	D	R	D	G	N	G	K	L	G	L	V	E
	D	L	D	K	S	G	S	M		A	Y	E
Fetuins												
human	E	G	D	C	D	F	Q	L	L	K	L	D
pig	E	G	D	C	D	F	H	V	L	K	Q	D
sheep	E	G	D	C	D	I	H	V	L	K	Q	D
cow	E	G	D	C	D	I	H	V	L	K	Q	D
rat	E	G	D	C	D	F	H	I	L	K	Q	D
CONSENSUS:	D	x	D N	x	D N S	G	x	I L V	G S T D N E	x	x	E

1990), thus making domain D1 a likely candidate for the Ca^{2+} binding site postulated by Wilson *et al.,* (1977). Domain D3 contains an aromatic-rich and strongly-conserved sequence (residues 292 to 302, sheep sequence, see Fig. 24), making this also a potential metal binding domain. However, direct experimental evidence is needed to define more precisely the role of domain D1 (or possibly other domains) in the binding of divalent cations.

Unlike the cystatin-like domains D1 and D2, the terminal domain D3 has not, as yet, been identified in any mammalians proteins, other than the fetuins. Indeed, no sequences of significant homology to this D3 region were found on database searching (data not shown). Domain D3 also contains the "connecting peptide" region (Lee *et al.,* 1987a). This 40 amino acid sequence had not previously been seen in the protein sequences of the "A" (Yoshioka *et al.,* 1986) or "B" chains (Gejyo *et al.,* 1983) of α_2-HS glycoprotein purified from human plasma. However, more recently Kellermann *et al.,* (1989) reported the purification of a "heavy" chain of human α_2-HS glycoprotein, which contained at its carboxyl terminus 39 of the 40 amino acids of the connecting peptide, suggesting that the second proteolytic cut to release the peptide was an artefact of storage and/or purification (see Fig. 7). A very similar mechanism to generate two disulphide-linked chains has been reported for human haptoglobin (Yang *et al.,* 1983) and tissue-type plasminogen activator (Pennica *et al.,* 1983). A further notable feature of domain D3 is the presence of a region unusually rich in proline, valine and alanine residues; indeed the three amino acids account for more than 60% of the residues in the D3a region of sheep and pig fetuin. Subdomain D3a contains a hydrophobic core region of 28 to 32 amino acid residues. Although this structural element is highly variable between the members of the fetuin family its overall hydrophobicity is well conserved. No rôle has been attributed to this hydrophobic region, but it does provide a primary candidate for the postulated lipid binding of mammalian fetuins (Kumbla *et al.,* 1989; Cayatte *et al.,* 1990). See also the section on O-glycosylation sites.

Conclusions.

The results presented and discussed in this thesis firmly establish the fetuins as a novel family within the cystatin superfamily. The fetuin proteins are made in the liver, comprise 3 domains (two homologous cystatin-like repeats at the amino-terminus and a unique carboxyl-terminal domain) and are extensively post-translationally modified. The protein is widely distributed, not confined to animals of the order *Artiodactyla* as once thought. The human plasma protein α_2-HS glycoprotein is the homologue of bovine fetuin. The protein encoded by cDNA clone pp63 is rat fetuin. The proposed domainal organisation of the protein is exactly reflected in the structure of the fetuin gene, 3 exons encoding each cystatin domain, as in other proteins in the cystatin superfamily, the unique carboxyl-terminal domain being encoded by a single exon. Clearly further studies are needed to define functionally relevant portions of the fetuin molecules. In particular, proteolytic dissection of the fetuins, into their constituent domains to allow functional analysis of the isolated modules will aid in further delineating the structural basis of fetuin function(s). Moreover, now that six fetuins have been cloned, recombinant strategies using truncated and modified versions of the proteins will provide useful probes to aid structure / function studies. Deletion and reporter gene analysis of the fetuin gene promoter region will be required to fully characterise the potential binding sites identified and to establish the molecular basis of fetuin's developmentally-regulated and tissue-specific expression and its expression in the acute phase response. *In situ* hybridisation experiments will be required to finally answer the question as to whether or not fetuin is made in the developing brain and given the cDNA sequences, probes can now be made. Finally, it remains to be seen whether the protein encoded by the clone pp63 genuinely corresponds to the phosphorylated glycoprotein PP^{63}; that is whether the fetuins are insulin receptor tyrosine kinase inhibitors, as PP^{63} was reported to be.

Future experiments.

In common with most scientific endeavour there is no "end" to this study; there is a large number of unanswered questions and, doubtless, in answering these, many others will arise.

In no particular order, there are three broad headings under which each of the questions fits more-or-less well; 1) the tissue-specific and developmentally-regulated expression of the fetuins, 2) the detailed physical structure of the proteins and 3) their function(s).

Tissue-specific and developmentally-regulated expression of fetuin.

The most obvious question remains whether or not *in situ* synthesis accounts for or contributes to the observed presence of fetuin in the developing brain. As discussed above, by the methods used in this study, it was not possible to detect fetuin mRNA in fetal sheep brain tissue at any of the ages examined. One way to tackle this is *in situ* hybridisation on tissue sections which immediately overcomes the dilution problem mentioned above. Now that the rat fetuin gene has been cloned and its promoter region sequenced, a further way of studying this has become available. It is now possible to hook up a reporter gene (the *E. coli lac Z* gene) to the rat fetuin promoter, to make transgenic mice with the construct, and to study, histochemically using X-gal, the distribution of the reporter gene product (the enzyme β-galactosidase). Furthermore, if expression is observed in the brain (or any other tissue), deletion analysis of the same construct will enable the identification of the *cis*-elements responsible.

The study of the developmental expression of the fetuins will have, in part, at least, to be species-specific. The levels of the protein plasma varies markedly with species; in pig, the plasma

concentration rises throughout gestation; in sheep, the level is constant throughout gestation, and in both, the fetal concentrations are an order of magnitude greater than those in the adult animal. In man, however, from the limited number of reports in the literature, human α_2-HS glycoprotein is essentially not a fetal protein, being present at approximately the same level in fetal and adult plasma.

Detailed physical structure of the proteins.

The recent publication of the X-ray crystal structure of human stefin B complexed with a cysteine protease and the known crystal structure of chicken cystatin will, perhaps, enable some progress to be made on a structural front. Given the sequence homology, and the knowledge that even the sequence differences must be confined within the same disulphide loop structure, it may be possible to model the likely structures of the cystatin domains the fetuins on the basis of the stefin and cystatin structures. If nothing else, it may provide a definitive answer to the question as to whether or not the cystatin domains in fetuin (and histidine rich glycoprotein) are viable cysteine protease inhibitors. Recent work by Prof W. Müller-Esterl and colleagues using proteins purified by Dr. K.M. Dziegielewska, suggests that at best fetuin is a very weak inhibitor of papain or cathepsins.

Such modelling is clearly no replacement for a proper X-ray crystal structure of fetuin itself. After all, only "two thirds" of the fetuin molecule is cystatin-like. The carboxyl-terminal domain bears no significant homology to any other protein currently in the sequence databases.

The function(s) of fetuin.

The function of the fetuin proteins is the one field in which there has been little progress in recent years. Indeed, it is probably

fair to argue that this study has even been reversed, given the number of recent reports in which it has been shown that previously documented properties of fetuin are due to copurifying contaminant proteins. Clearly, more detailed analyses are required to assess the purity of fetuin preparations before studying their effects on any system. Furthermore many of the literature reports must now be considered with a degree of scepticism, given the purification methods used.

The one *in vivo* property that has been studied is that of the acute phase response *(vide supra).* However, any hypothesis as to the true rôle of fetuin in this must explain why in rat and man fetuin is a negative acute phase protein, while in cattle it is a positive one. A similar point can be made with respect to the levels of fetuin in plasma during development. The absolute levels, and indeed the changes in levels, seen through development are quite different in human, pig and sheep.

Intuitively, the function of any protein present in plasma at over 1000mg/100ml (as fetuin is in fetal cattle and pigs) cannot be very subtle! The most obvious "bulk property" is that of a "carrier", but, again, any claim to have found the function in cattle, sheep or pigs must also explain why in man, the concentration of α_2-HS glycoprotein in fetal plasma is less than 100mg/100ml; presumably other plasma proteins in the human fetus contribute to or carry out a similar rôle?

Clearly, much more work is required on a number of different fronts to explain the function(s) of the plasma glycoprotein fetuin. The collections of clones and constructs described in this thesis should provide the foundation for future molecular biological studies.

****

Appendices.

Appendix I.

The chromosomal localisation of human plasma proteins.

This table is modified from that presented by Bowman & Yang, (1987) and includes a number of further localisations reported since then. See also Moore *et al.*, (1985).

protein	symbol	chromosome	arm	region	reference(s)
serum amyloid P component	APCS	1	q	q12-q23	Mantzouranis et al., (1985)
apolipoprotein A-II	APOA2	1		p21-qter	Lackner et al., (1984); Rogne et al., (1989)
complement component 1q.β	C1QB	1	p		Solomon et al., (1986)
complement component 8A	C8A	1	p	p36.2-p22.1	Alper et al., (1983); Kaufman et al., (1989)
complement component 8B	C8B	1			Rogde et al., (1986)
coagulation factor III	F3	1	p	pter-p21	Carson et al., (1985)
C-reactive protein	CRP	1	q	q12-q23	Whitehead et al., (1983)
apolipoprotein B	APOB	2	p	pter-p23	Law et al., (1985)
immunoglobulin polypeptide *k*	IGK	2	p	p12	Malcolm et al., (1982)
protein C	PROC	2			Rocchi et al., (1986)
α_2-HS glycoprotein	AHSG	3	q	q27-q29	Eiberg et al., (1984); Cox & Francke, (1985); Arnaud et al., (1987); Magnuson et al., (1988)
ceruloplasmin	CP	3	q	q25	Weitkamp, (1983); Yang et al., (1986)
transferrin	TF	3	q	q21-q24	Yang et al., (1984); Moore et al., (1985)

protein	symbol	chromosome	arm	region	reference(s)
albumin	ALB	4	q	q11-q13	Kurnit et al., (1982)
α-fetoprotein	AFP	4	q	q11-q13	Harper & Dugaiczyk, (1983); Murray et al., (1985)
complement factor I	IF	4	q	q25	Shiang et al., (1989)
fibrinogen α or A chain	FGA	4	q	q26-q28	Henry et al., (1984); Kant et al., (1985)
fibrinogen β or B chain	FGB	4	q	q26-q28	Henry et al., (1984); Kant et al., (1985)
fibrinogen Y chain	FGG	4	q	q26-q28	Kant et al., (1985)
Group-specific component, Gc	GC	4	q	q12-q13	Weitkamp et al., (1970); Yang et al., (1985); Moore et al., (1985); McCombs et al., (1986)
complement component C9	C9	5			Rogne et al., (1989); Abbott et al., (1989)
properdin factor B	BF	6	p	p21.3	Allen, (1974); Olaisen et al., (1983)
complement component 2	C2	6	p	p21.3	Alper, (1981); Carroll et al., (1984)
complement component 4A	C4A	6	p	p21.3	Alper, (1981); Carroll et al., (1984)
complement component 4B	C4B	6	p	p21.3	Alper, (1981); Carroll et al., (1984)
coagulation factor XII	F12	6	p	p23	Pearson et al., (1982)
coagulation factor XIII A	F13A	6	p	p24-p25	Olaisen et al., (1985); Board et al., (1988)
plasminogen	PLG	6	q	q25-qter	Murray et al., (1985)
ceruloplasmin, processed pseudogene	CPP	8	q	q21.13-q231	Wang et al., (1988)

protein	symbol	chromosome	arm	region	reference(s)
orosomucoid (α_1-acid glycoprotein)	ORM	9	q	q31-q34.1	Cox & Francke, (1985); Webb et al., (1988)
apolipoprotein A-I	APOA1	11	q	q13-qter	Cheung et al., (1984)
apolipoprotein A-IV	APOA4	11	q	q13-qter	Schamaun et al., (1984)
apolipoprotein C-III	APOC3	11	q	q13-qter	Karathanasis et al., (1983)
ferritin H	FTH	11	q	q13	Boyd et al., (1984)
α_2-macroglobulin	A2M	12	p	p12.3-p13.3	Kan et al., (1985); Devriendt et al., (1989); Fukushima et al., (1988)
immunoglobulin polypeptides;					
α_1	IGHA1	14	q	q32.3	Croce et al., (1979); McBride et al., (1982)
α_2	IGHA2	14	q	q32.3	Croce et al., (1979); McBride et al., (1982)
δ	IGHD	14	q	q32.3	Croce et al., (1979); McBride et al., (1982)
ϵ	IGHE	14	q	q32.3	McBride et al., (1982)

protein	symbol	chromosome	arm	region	reference(s)
immunoglobulin polypeptides;					
Y 1,2,3,4	IGHG 1,2,3,4	14	q	q32.3	Hobart et al., (1981)
μ	IGHM	14	q	q32.3	Croce et al., (1979); Hobart et al., (1981)
heavy	IGHV	14	q	q32.3	Hobart et al., (1981); McBride et al., (1982)
α_1-antitrypsin (protease inhibitor)	PI	14	q	q32.1	Darlington et al., (1982); Pearson et al., (1982); Lai et al., (1983)
β_2-microglobulin	B2M	15	q	q22	Goodfellow et al., (1975)
haptoglobin	HP	16	q	q22.1	McGill et al., (1984)
haptoglobin-related	HPR	16	q	q22.1	Maeda et al., (1984)
transthyretin (prealbumin)	PALB	18	q	q11.2-q12.1	Wallace et al., (1985)
α_2-plasmin inihibitor (α_2-antiplasmin)	PLI	18	-	p11.1-q11.2	Kato et al., (1988)
apolipoprotein C-I	APOC1	19		cen-q13.2	Tata et al., (1985)
apolipoprotein C-II	APOC2	19		cen-q13.2	Humphries et al., (1984)
apolipoprotein E	APOE	19		cen-q13.2	Olaisen et al., (1982)

protein	symbol	chromosome	arm	region	reference(s)
complement component 3	C3	19	p	p13.3-p13.2	Lachmann, (1982)
ferritin L	FTL	19	q	q13.3-q13.4	Caskey et al., (1983); McGill et al., (1984)
immunoglobulin polypeptides;					
λ-constant	IGLC	22	q	q11.1-q11.2	Erikson et al., (1981)
λ-variable	IGLV	22	q	q11.1-q11.2	Anderson et al., (1984)
coagulation factor XIII C	F8C	X	q	q28	Gitshier et al., (1985)
coagulation factor IX	F9	X	q	q26-q27.3	Camerino et al., (1984)
thyroxine binding globulin	TBG	X			Daiger et al., (1982)

Appendix II.

AHSG allelic frequencies found in population studies.

Number of subjects and origin		AHSG*1	AHSG*2	AHSG*3	AHSG*4	AHSG*5	AHSG*B	AHSG*7	AHSG*10	AHSG*11	AHSG*16	AHSG*17	AHSG*18	Reference
[a]	North America	0.64	0.36											Anderson & Anderson, (1979)
68	Canada	0.64	0.36	0.015										Cox & Andrews, (1983)
166	Germany	0.654	0.322	0.015	0.009									Weidinger et al., (1984)
150	US caucasians	0.653	0.346											Boutin et al., (1985)
2050	Japan	0.7356	0.2639			0.0005								Umetsu et al., (1984b)
1003	Japan	0.7338	0.2662											Umetsu et al., (1984a)
300	Izumo, Japan	0.7233	0.2767											Yuasa et al., (1985)
400	Yamaguchi, Japan	0.7325	0.2675											Yuasa et al., (1985)
397	Okinawa, Japan	0.7670	0.2065			0.0264								Yuasa et al., (1985)
140	Nepal	0.7571	0.2429											Yuasa et al., (1985)
215	caucasians	0.6419	0.3535	0.0046										Cox et al., (1986)
71	blacks	0.6901	0.2606				0.0493							Cox et al., (1986)
115	Philippines	0.6870	0.3130											Umetsu et al., (1988)
382	white Europeans	0.6466	0.3469		0.0052				0.0013					Westwood et al., (1987b)
205	Asia	0.8073	0.1878		0.0049									Westwood et al., (1987b)
119	Afro-Carribean	0.6597	0.2353						0.0966	0.0084				Westwood et al., (1987a,b)
1074	Fukuoka, Japan	0.7165	0.2817			0.0009					0.0005		0.0005	Fukuma et al., (1990)
514	Tsushima, Japan	0.6916	0.3045			0.0010		0.0019				0.0010		Fukuma et al., (1990)
103	Koshiki, Japan	0.7282	0.2718											Fukuma et al., (1990)
110	Libya	0.8364	0.1636											Sebetan & Heshmat, (1988)
1000	Brisbane	0.6420	0.3530	0.0050										Thomas, (1989a)
256	Fukui, Japan	0.7637	0.2363											Kishi et al., (1988)

[a] number of subjects tested not indicated.

Literature Cited.

Abbott, C., West, L., Povey, S., Jeremiah, S., Murad, Z., DiScipio, R., & Fey, G. (1989) The gene for human complement component C9 mapped to chromosome 5 by polymerase chain reaction. Genomics 4, 606-609

Abdullah, M., Crowell, J.A., Tres, L.L., & Kierszenbaum, A.L. (1986) Fetuin: a serum component associated with rat Sertoli and spermatogenic cells in coculture. J. Cell. Physiol. 127, 463-472

Abe, K., Emori, Y., Kondo, H., Suzuki, K., & Arai, S. (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. J. Biol. Chem. 262, 16793-16797

Abiodun, P.O., & Olomu, I.N. (1987) Alpha$_2$ HS-glycoprotein levels in children with protein-energy malnutrition and infections. J. Pediatr. Gastroenterol. Nutr. 6, 271-275

Abiodun, P.O., Ihongbe, J.C., & Dati, F. (1985) Decreased levels of alpha$_2$ HS-glycoprotein in children with protein-energy-malnutrition. Eur. J. Pediatr. 144, 368-369

Abrahamson, M., Olafsson, I., Palsdottir, A., Ulvsbäck, M., Lundwall, A., Jensson, O. & Grubb, A. (1990) Structure and expression of the human cystatin C gene. Biochem. J. 268, 287-294

Adolf, G.R., Maurer-Fogy, I., Kalsner, I., & Cantell, K. (1990) Purification and characterization of natural human interferon ω1. Two alternative cleavage sites for the signal peptidase. J. Biol. Chem. 265, 9290-9295

Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T., & Kishimoto, T. (1990) A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J. 9, 1897-1906

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., & Watson, J.D. (1983) Molecular biology of the cell. Garland Publishing Inc., New York.

Alcaraz, G., Marti, J., Moinier, D., & Fougereau, M. (1981) NH_2-terminal sequence of calf fetuin. Biochem. Biophys. Res. Commun. 99, 30-36

Alexander, F., Young, P.R., & Tilghman, S.M. (1984) Evolution of the albumin:α-fetoprotein ancestral gene from the amplification of a 27 nucleotide sequence. J. Mol. Biol. 173, 159-176

Al-Hashimi, I., Dickinson, D.P., & Levine, M.J. (1988) Purification, molecular cloning, and sequencing of salivary cystatin SA-I. J. Biol. Chem. 263, 9381-9387

Ali M., Mujoo K. & Sahib M. K. (1983) Synthesis and secretion of alpha-fetoprotein and albumin by newborn rat brain cells in culture. Dev. Brain Res. 6, 47-55

Allen, F.H. Jr. (1974) Linkage of HL-A and GBG. Vox. Sang. 27, 382-384

Alper, C.A., Marcus, D., Raum, D., Petersen, B.H., & Spira, T.J. (1983) Genetic polymorphism in C8 β-chains. Evidence for two unlinked genetic loci for the eighth component of human complement (C8). J. Clin. Invest. 72, 1526-1531

Alper, C.A. (1981) Complement and the MHC. In "The role of the major histocompatibility complex in immunobiology" Ed. Dorf, M.E., Garland STPM Press, New York.

Anderson, L., & Anderson, N.G. (1977) High resolution two-dimensional electrophoresis of human plasma proteins. Proc. Natl. Acad. Sci. USA 74, 5421-5425

Anderson, N.L., & Anderson, N.G. (1979) Microheterogeneity of serum transferrin, haptoglobin and α_2HS glycoprotein examined by high resolution two-dimensional electrophoresis. Biochem. Biophys. Res. Commun. 88, 258-265

Anderson, K.P., Croyle, M.L., & Lingrel, J.B. (1989) Primary structure of a gene encoding rat T-kininogen. Gene 81, 119-128

Anderson, M.L.M., Szajnert, M.F., Kaplan, J.C., McColl, L., & Young, B.D. (1984) The isolation of a human Ig V_λ gene from a recombinant library of chromosome 22 and estimation of its copy number. Nucleic Acids Res. 12, 6647-6661

Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Rahmsdorf, H.J., Jonat, C., Herrlich, P., & Karin, M. (1987) Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. Cell 49, 729-739

Araki, T., Yoshioka, Y., & Schmid, K. (1989) The position of the disulfide bonds in human plasma α_2HS-glycoprotein and the repeating double disulfide bonds in the domain structure. Biochim. Biophys. Acta 994, 195-199

Arcone, R., Gualandi, G., & Ciliberto, G. (1988) Identification of sequences responsible for acute-phase induction of human C-reactive protein. Nucleic Acids Res. 16, 3195-3207

Arnaud, P., Emerson, D.L., & Gianazza, E. (1983) Interaction of α_2-HS-glycoprotein with immobilized triazine dyes. Biochim. Biophy. Acta 749, 270-275

Arnaud, P., Mietz, J.A., Grossman, Z., & McBride, O.W. (1987) Isolation and characterization of a cDNA clone for human alpha$_2$-HS glycoprotein. Protides Biol. Fluids 35, 135-138

Arnaud, P., Miribel, L., & Emerson, D.L. (1988) α_2-HS glycoprotein. Methods Enzymol. 163, 431-441

Ashburner, M., & Bonner, J.J. (1979) The induction of gene activity in Drosophila by heat shock. Cell 17, 241-254

Ashton, B.A., Höhling, H.-J., & Triffitt, J.T. (1976) Plasma proteins present in human cortical bone: enrichment of the α_2HS-glycoprotein. Calcif. Tiss. Res. 22, 27-33

Ashton, B.A., & Smith, R. (1980) Plasma α_2HS-glycoprotein concentration in Paget's disease of bone: its possible significance. Clin. Sci. 58, 435-438

Ashton, B.A., Triffitt, J.T., & Herring, G.M. (1974) Isolation and partial characterization of a glycoprotein from bovine cortical bone. Eur. J. Biochem. 45, 525-533

Auberger, P., Falquerho, L., Contreres, J.O., Pages, G., Le Cam, G., Rossi, B., & Le Cam, A. (1989) Characterization of a natural inhibitor of the insulin receptor tyrosine kinase. cDNA cloning, purification, and anti-mitogenic activity. Cell 58, 631-640

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.M., Seidman, J.G., Smith, J.A., Struhl, K., Wang-Iverson, P., & Bonitz, S.G. (1989) Short protocols in molecular biology. A compendium of methods from 'Current protocols in molecular biology' John Wiley & Sons, New York.

Baenziger, J.U., & Fiete, D. (1979) Structure of the complex oligosaccharides of fetuin. J. Biol. Chem. 254, 789-795

Bairoch, A., & Cox, J.A. (1990) EF-hand motifs in inositol phospholipid-specific phospholipase C. FEBS Letts. 269, 454-456

Barboriak, J.J., Meschia, G., Barron, D.H., & Cowgill, G.R. (1958) Blood plasma proteins in fetal goats and sheep. Proc. Soc. Exptl. Biol. 99, 635-637

Barnes, D., & Sato, G. (1980) Serum-free cell culture: a unifying approach. Cell 22, 649-655

Barrett, A.J. (1981) Cystatin, the egg white inhibitor of cysteine proteinases. Methods Enzymol. 80, 771-778

Barrett, A.J. (1986) The cystatins: a diverse superfamily of cysteine peptidase inhibitors. Biomed. Biochim. Acta 45, 1363-1374

Barrett, A.J. (1987) The cystatins: a new class of peptidase inhibitors. Trends Biochem. Sci. 12, 193-196

Baskies, A.M., Chretien, P.B., Weiss, J.F., Makuch, R.W., Beveridge, R.A., Catalona, W.J., & Spiegel, H.E. (1980) Serum glycoproteins in cancer patients: first report of correlations with *in vitro* and *in vivo* parameters of cellular immunity. Cancer 45, 3050-3060

Baumann, H., Richards, C., & Gauldie, J. (1987) Interaction among hepatocyte-stimulating factors, interleukin 1 and glucocorticoids for regulation of acute phase plasma proteins in human hepatoma (HepG2) cells. J. Immunol. 139, 4122-4128

Baumhueter, S., Mendel, D.B., Conley, P.B., Kuo, C.J., Turk, C., Graves, M.K., Edwards, C.A., Courtois, G., & Crabtree, G.R. (1990) HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF. Genes Dev. 4, 372-379

Becker-Andrē, M., & Hahlbrock, K. (1989) Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a *P*CR *a*ided *t*ranscript *t*itration assa*y* (PATTY). Nucleic Acids Res. 17, 9437-9446

Begbie R. (1974) Studies on fetuin from foetal bovine serum. The composition and amino acid sequences of glycopeptides from fetuin. Biochim. Biophys. Acta 371, 549-567

Bendiak, B., Harris-Brandts, M., Michnick, S.W., Carver, J.P., & Cumming, D.A. (1989) Separation of the complex asparagine-linked oligosaccharides of the glycoprotein fetuin and elucidation of three triantennary structures having sialic acids linked only to galactose residues. Biochemistry 28, 6491-6499

Benno, R.H., & Williams, T.H. (1978) Evidence for intracellular localization of alpha-fetoprotein in the developing rat brain. Brain Res. 142, 182-186

van den Berg, E.A., le Clercq, E., Kluft, C., Koide, T., van der Zee, A., Oldenburg, M., Wijnen, J.T., & Khan, P.M. (1990) Assignment of the human gene for histidine-rich glycoprotein to chromosome 3. Genomics 7, 276-279

Bergh, M.L.E., Hooghwinkel, G.J.M., van den Eijnden, D.H. (1983) Biosynthesis of the *O*-glycosidically linked oligosaccharide chains of fetuin. Indications for an α-*N*-acetylgalactosaminide α2→6 sialyltransferase with a narrow acceptor specificity in fetal calf liver. J. Biol. Chem. 258, 7430-7436

Bergmann, F.H., Levine, L., & Spiro, R.G. (1962) Fetuin: immunochemistry and quantitative estimation in serum. Biochim. Biophys. Acta 58, 41-51

Berman, E. (1986) Reinvestigation of the carbohydrate chains of calf fetuin using ^{13}C-N.M.R. spectroscopy. Carbohydrate Res. 152, 33-46

Berman E. & Bendel P. (1986) One- and two-dimensional 90.5-MHz ^{13}C-NMR spectroscopy of the *N*-linked triantennary oligosaccharide units of calf fetuin. FEBS Letts. 204, 257-260

Berman, E., Dabrowski, U., & Dabrowski, J. (1988) A two-dimensional ^{1}H-N.M.R. (500 MHz) and ^{13}C-N.M.R. (125 MHz) study of *N*-linked glycopeptides derived from calf fetuin. Carbohydrate Res. 176, 1-15

Billiau, A. (1986) BSF-2 is not just a differentiation factor. Nature (London) 324, 415

Billiau, A. (1987) Interferon β_2 as a promoter of growth and differentiation of B cells. Immunol. Today 8, 84-87

Birnboim, H.C., & Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513-1523

Bloch, B., Popovici, T., Levin, M.J., Tuil, D., & Kahn, A. (1985) Transferrin gene expression visualized in oligodendrocytes of the rat brain by using *in situ* hybridization and immunohistochemistry. Proc. Natl. Acad. Sci. USA 82, 6706-6710

Board, P.G., Webb, G.C., McKee, J., & Ichinose, A. (1988) Localization of the coagulation factor XIII A subunit gene (F13A) to chromosome bands 6p24→p25. Cytogenet. Cell Genet. 48, 25-27

Bohmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G., & Mattaj, I.W. (1987) A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature (London) 325, 268-272

Borgmeyer, U., Nowock, J., & Sippel, A.E. (1984) The TGGCA-binding protein: a eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. Nucleic Acids Res. 12, 4295-4311

Boutin, B., Feng, S.H., & Arnaud, P. (1985) The genetic polymorphism of alpha$_2$-HS glycoprotein: study by ultrathin-layer isoelectric focusing and immunoblot. Am. J. Hum. Genet. 37, 1098-1105

Bowman, B.H., & Yang, F. (1987) DNA sequencing and chromosomal locations of human plasma protein genes. In "The Plasma Proteins. Vol. 5. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Bowman, B.H., Yang, F., & Adrian, G.S. (1990) Expression of human plasma protein genes in ageing transgenic mice. Bioessays 12, 317-322

Boyd, D., Jain, S.K., Crampton, J., Barrett, K.J., & Drysdale, J. (1984) Isolation and characterization of a cDNA clone for human ferritin heavy chain. Proc. Natl. Acad. Sci. USA 81, 4751-4755

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254

Broadwell, R.D., & Brightman, M.W. (1976) Entry of peroxidase into neurons of the central and peripheral nervous systems from extracerebral and cerebral blood. J. Comp. Neurol. 166, 257-284

Brown, J.P., Hewick, R.M., Hellström, I., Hellström, K.E., Doolittle, R.F., & Dreyer, W.J. (1982) Human melanoma-associated antigen p97 is structurally and functionally related to transferrin. Nature (London) 296, 171-173

Brown, W.M., & Reynolds, M.L. (1988) An autoradiographic study to determine the origin of α_2-HS-positive cells in the developing forebrain of the rat. J. Physiol. 400, 65p

Brown, W.M., Christie, D.L., Dziegielewska, K.M., Saunders, N.R., & Yang, F. (1991a) The rat protein encoded by clone "pp63" is a fetuin / α_2-HS glycoprotein-like molecule, but is it the tyrosine kinase inhibitor PP^{63}? Cell, in press.

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., & Saunders, N.R. (1989a) Nucleotide and deduced amino acid sequence of a γ subunit of bovine fibrinogen. Nucleic Acids Res. 17, 6397

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., & Saunders, N.R. (1989b) Nucleotide and deduced amino acid sequence of sheep serum albumin. Nucleic Acids Res. 17, 10495

Brown, W.M., Dziegielewska, K.M., Foreman, R.C., Saunders, N.R., & Wu, Y. (1989c) Nucleotide and deduced amino acid sequence of sheep α1 antitrypsin. Nucleic Acids Res. 17, 6398

Brown, W.M., Dziegielewska, K.M., Saunders, N.R., Christie, D.L., Nawratil, P., & Müller-Esterl, W. (1991b) The nucleotide and deduced amino acid structures of sheep and pig fetuin. Common structural features of the mammalian fetuin family. Eur. J. Biochem., submitted

Buchman, A.R., Burnett, L., & Berg, P. (1980) The SV40 nucleotide sequence. In "Molecular biology of tumor viruses, part II" Ed. Tooze, J., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. pp. 799-829

Butler, E.T., & Chamberlin, M.J. (1982) Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. J. Biol. Chem. 257, 5772-5778

Camerino, G., Grzeschik, K.H., Jaye, M., de la Salle, H., Tolstoshev, P., Lecocq, J.P., Heilig, R., & Mandel, J.L. (1984) Regional localization on the human X chromosome and polymorphism of the coagulation factor IX gene (hemophilia B locus). Proc. Natl. Acad. Sci. USA 81, 498-502

Campbell, R.D., Bentley, D.R., Carroll, M.C., & Morley, B.J. (1985) Human complement genes of the major histocompatibility complex. Protides Biol. Fluids 35, 43-46

Carroll, M.C., Campbell, R.D., Bentley, D.R., & Porter, R.R. (1984) A molecular map of the human major histocompatibility complex class III region linking complement genes C4, C2 and factor B. Nature (London) 307, 237-241

Carson, S.D., Henry, W.M., & Shows, T.B. (1985) Tissue factor gene localized to human chromosome 1 (1pter→1p21). Science 229, 991-993

Caruthers, M.H. (1985) Gene synthesis machines: DNA chemistry and its uses. Science 230, 281-285

Caskey, J.H., Jones, C., Miller, Y.E., & Seligman, P.A. (1983) Human ferritin gene is assigned to chromosome 19. Proc. Natl. Acad. Sci. USA 80, 482-486

Cavanagh, M.E., & Warren, A. (1985) The distribution of native albumin and foreign albumin injected into lateral ventricles of prenatal and neotnatal rat forebrains. Anat. Embryol. 17, 345-351

Cayatte, A.J., Kumbla, L., & Subbiah, M.T.R. (1990) Marked acceleration of exogenous fatty acid incorporation into cellular triglycerides by fetuin. J. Biol. Chem. 265, 5883-5888

Cereghini, S., Raymondjean, M., Carranca, A.G., Herbomel, P., & Yaniv, M. (1987) Factors involved in control of tissue-specific expression of albumin gene. Cell 50, 627-638

Cheung, P., Kao, F.-T., Law, M.L., Jones, C., Puck, T.T., & Chan, L. (1984) Localization of the structural gene for human apolipoprotein A-1 on the long arm of human chromosome 11. Proc. Natl. Acad. Sci. USA 81, 508-511

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., & Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18, 5294-5299

Chomczynski, P., & Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159

Christie, D.L., Dziegielewska, K.M., Hill, R.M., & Saunders, N.R. (1987) Fetuin: the bovine homologue of human α_2HS glycoprotein. FEBS Letts. 214, 45-49

Ciliberto, G., Dente, L., & Cortese, R. (1985) Cell-specific expression of a transfected human α_1-antitrypsin gene. Cell 41, 531-540

Clausen, P.P., & Thomsen, P. (1978) Demonstration of hepatitis B-surface antigen in liver biopsies. Acta path. microbiol. scand. Sect. A 86, 383-388

Cochran, B.H., Zullo, J., Verma, I.M., & Stiles, C.D. (1984) Expression of the c-*fos* gene and of an *fos*-related gene is stimulated by platelet-derived growth factor. Science 226, 1080-1082

Colamonici, O.R., Trepel, J.B., Vidal, C.A., & Neckers, L.M. (1986) Phorbol ester induces c-*sis* gene transcription in stem cell line K-562. Mol. Cell. Biol. 6, 1847-1850

Colclasure, G.C., Lloyd, W.S., Lamkin, M., Gonnerman, W., Troxler, R.F., Offner, G.D., Bürgi, W., Schmid, K., & Nimberg, R.B. (1988) Human serum α_2HS-glycoprotein modulates *in vitro* bone resorption. J. Clin. Endocrin. Metab. 66, 187-192

Cole, T., Inglis, A.S., Roxburgh, C.M., Howlett, G.J., & Schreiber, G. (1985) Major acute phase α_1-protein of the rat is homologous to bovine kininogen and contains the sequence for bradykinin: its synthesis is regulated at the mRNA level. FEBS Letts. 182, 57-61

Confavreux, C., Gianazza, E., Chazot, G., Lasne, Y., & Arnaud, P. (1982) Silver stain after isoelectric focusing of unconcentrated cerebrospinal fluid: visualization of total protein and direct immunofixation of immunoglobulin G. Electrophoresis 3, 206-210

Costa, R.H., Grayson, D.R., & Darnell J.E. Jr. (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α1-antitrypsin genes. Mol. Cell. Biol. 9, 1415-1425

Costa, R.H., Lai, E., & Darnell, J.E. Jr. (1986) Transcriptional control of the mouse prealbumin (transthyretin) gene: both promoter sequences and a distinct enhancer are cell specific. Mol. Cell. Biol. 6, 4697-4708

Courtois, G., Baumhueter, S., & Crabtree, G.R. (1988) Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. Proc. Natl. Acad. Sci. USA 85, 7937-7941

Courtois, G., Morgan, J.G., Campbell, L.A., Fourel, G., & Crabtree, G.R. (1987) Interaction of a liver-specific nuclear factor with the fibrinogen and α_1-antitrypsin promoters. Science 238, 688-692

Courtoy, P.J., Lombart, C., Feldmann, G., Moguilevsky, N., & Rogier, E. (1981) Synchronous increase of four acute phase proteins synthesized by the same hepatocytes during the inflammatory reaction. A combined biochemical and morphologic kinetics study in the rat. Lab. Invest. 44, 105-115

Cox, D.W., & Andrews, B.J. (1983) Silver stain immunofixation for α_2HS-glycoprotein: a new method for detection of protein heterogeneity. In "Electrophoresis '82", Walter de Gruyter & Co., New York. pp. 243-247

Cox, D.W., & Francke, U. (1985) Direct assignment of orosomucoid to human chromosome 9 and α2HS-glycoprotein to chromosome 3 using human fetal liver x rat hepatoma hybrids. Hum. Genet. 70, 109-115

Cox, D.W., Andrews, B.J., & Wills, D.E. (1986) Genetic polymorphism of α_2HS-glycoprotein. Am. J. Hum. Genet. 38, 699-706

Crawford, S.M. (1984) Alpha-2-HS glycoprotein in the hypercalcaemia of multiple myeloma. Br. J. Cancer 49, 813-815

Croce, C.M., Shander, M., Martinis, J., Cicurel, L., D'Ancona, G.G., Dolby, T.W., & Koprowski, H. (1979) Chromosomal location of the genes for human immunoglobulin heavy chains. Proc. Natl. Acad. Sci. USA 76, 3416-3419

Cumming, D.A., Hellerqvist, C.G., Harris-Brandts, M., Michnick, S.W., Carver, J.P., & Bendiak, B. (1989) Structures of asparagine-linked oligosaccharides of the glycoprotein fetuin having sialic acids linked to *N*-acetylglucosamine. Biochemistry 28, 6500-6512

Dagert, M., & Ehrlich, S.D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene 6, 23-28

Daiger, S.P., Rummel, D.P., Wang, L., & Cavalli-Sforza, L.L. (1982) Detection of genetic variation with radioactive ligands. IV. X-linked, polymorphic genetic variation of thyroxin-binding globulin (TBG). Am. J. Hum. Genet. 33, 640-648

Darlington, G.J., Astrin, K.H., Muirhead, S.P., Desnick, R.J., & Smith, M. (1982) Assignment of human α_1-antitrypsin to chromosome 14 by somatic cell hybrid analysis. Proc. Natl. Acad. Sci. USA 79, 870-873

Darnell, J., Lodish, H., & Baltimore, D. (1986) Molecular Cell Biology. Revised Edition, Scientific American Books Inc., New York.

Darveau, A., Pelletier, J., & Sonenberg, N. (1985) Differential efficiencies of *in vitro* translation of mouse c-*myc* transcripts differing in the 5'-untranslated region. Proc. Natl. Acad. Sci. USA 82, 2315-2319

Daveau, M., Davrinche, C., Djelassi, N., Lemetayer, J., Julen, N., Hiron, M., Arnaud, P., & Lebreton, J.P. (1990) Partial hepatectomy and mediators of inflammation decrease the expression of liver α_2-HS glycoprotein gene in rats. FEBS Letts. 273, 79-81

Daveau, M., Davrinche, C., Julen, N., Hiron, M., Arnaud, P., & Lebreton, J.-P. (1988) The synthesis of human α-2-HS glycoprotein is down-regulated by cytokines in hepatoma HepG2 cells. FEBS Letts. 241, 191-194

Davis, D.L., Dibner, M.D., & Batley, J.E. (1986) Basic methods in molecular biology. Elsevier, Amsterdam.

Dayhoff, M.O., Barker, W.C., & Hunt, L.T. (1978) Protein superfamilies. In "Atlas of protein sequence and structure. Vol. 5, suppl. 3" Ed. Dayhoff, M.O., National Biomedical Research Foundation, Washington DC. pp. 1-8.

Denhardt, D.T. (1966) A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. 23, 641-646

Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M., & Darnell, J.E. Jr. (1981) Transcriptional control in the production of liver-specific mRNAs. Cell 23, 731-739

Devriendt, K., Zhang, J., van Leuven, F., van den Berghe, H., Cassiman, J.-J., & Marynen, P. (1989) A cluster of α_2-macroglobulin-related genes ($\alpha_2 M$) on human chromosome 12p: cloning of the pregnancy-zone protein gene and an $\alpha_2 M$ pseudogene. Gene 81, 325-334

Dickson, I.R., Bagga, M., & Paterson, C.R. (1983) Variations in the serum concentration and urine excretion of α_2HS-glycoprotein, a bone-related protein, in normal individuals and in patients with osteogenesis imperfacta. Calcif. Tiss. Int. 35, 16-20

Dickson, I.R., Poole, A.R., & Veis, A. (1975) Localisation of plasma α_2HS glycoprotein in mineralising human bone. Nature (London) 256, 430-432

Dickson, P.W., Aldred, A.R., Marley, P.D., Bannister, D., & Schreiber, G. (1986) Rat choroid plexus specializes in the synthesis and the secretion of transthyretin (prealbumin). Regulation of transthyretin synthesis in choroid plexus in independent from that in liver. J. Biol. Chem. 261, 3475-3478

Dickson, P.W., Aldred, A.R., Marley, P.D., Guo-Fen, T., Howlett, G.J., & Schreiber, G. (1985) High prealbumin and transferrin mRNA levels in the choroid plexus of rat brain. Biochem. Biophys. Res. Commun. 127, 890-895

Dugaiczyk, A., Ruffner, D.E., Minghetti, P.P., & Gibbs, P.E.M. (1985) Structure and evolution of the serum albumin gene family. Protides Biol. Fluids 35, 27-30

Dynan, W.S., & Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence specific DNA binding proteins. Nature (London) 316, 774-778

Dziegielewska, K.M. (1982) Proteins in fetal CSF and plasma. Ph.D. Thesis, University of London, UK.

Dziegielewska, K.M., & Saunders, N.R. (1988) The development of the blood-brain barrier: proteins in fetal and neonatal CSF, their nature, and origins. In "Handbook of Human Growth & Developmental Biology. Vol. 1: Neural, sensory, motor & integrative development. Part A Developmental Neurobiology". Eds. Meisami, E., & Timiras, P.S., CRC Press, Boca Raton, Florida.

Dziegielewska, K.M., Bock, E., Cornelis, M.E.P., Møllgård, K, New, H., & Saunders, N.R. (1983) Identification of fetuin in human and rat fetuses and in other species. Comp. Biochem. Physiol. 76A, 241-245

Dziegielewska, K.M., Brown, W.M., Casey, S.-J., Christie, D.L., Foreman, R.C., Hill, R.M., & Saunders, N.R. (1990) The complete cDNA and amino acid sequence of bovine fetuin. Its homology with α_2HS glycoprotein and relation to other members of the cystatin superfamily. J. Biol. Chem. 265, 4354-4357

Dziegielewska, K.M., Brown, W.M., Gould, C.C., Matthews, N., Sedgwick, J.E.C., & Saunders, N.R. (1991) Variations in the serum concentration of fetuin - the bovine equivalent of human α_2HS glycoprotein. J. Comp. Physiol., submitted

Dziegielewska. K.M., Evans, C.A.N., Fossan, G., Lorscheider, F.L., Malinowska, D.H., Møllgård, K., Reynolds, M.L., Saunders, N.R., & Wilkinson, S. (1980) Proteins in cerebrospinal fluid and plasma of fetal sheep during development. J. Physiol. 300, 441-455

Dziegielewska, K.M., Møllgård, K., Reynolds, M.L., & Saunders, N.R. (1987) A fetuin-related glycoprotein (α_2HS) in human embryonic and fetal development. Cell Tissue Res. 248, 33-41

Dziegielewska, K.M., Saunders, N.R., & Soreq, H. (1985) Messenger ribonucleic acid (mRNA) from developing rat cerebellum directs in vitro synthesis of plasma proteins. Dev. Brain Res. 23, 259-267

Dziegielewska, K.M., Saunders, N.R., Schejter, E.J., Zakut, H., Zevin-Sonkin, D., Zisling, R., & Soreq, H. (1986) Synthesis of plasma proteins in fetal, adult, and neoplastic human brain tissue. Dev. Biol. 115, 93-104

Edge, A.S.B., & Spiro, R.G. (1987) Presence of an *O*-glycosidically linked hexasaccharide in fetuin. J. Biol. Chem. 262, 16135-16141

Eiberg, H., Mohr, J., & Nielsen, L.S. (1984) A2HS: new methods of phenotyping and analysis of linkage relations: assignment to chromosome 3. Cytogenet. Cell Genet. 37, 461

Eisenberg, D. (1984) Three-dimensional structure of membrane and surface proteins. Ann. Rev. Biochem. 53, 595-623

Elder, P.K., Schmidt, L.J., Ono, T., & Getz, M.J. (1984) Specific stimulation of actin transcription by epidermal growth factor and cycloheximide. Proc. Natl. Acad. Sci. USA 81, 7476-7480

Elzanowski, A., Barker, W.C., Hunt, L.T., & Seibel-Ross, E. (1988) Cystatin domains in alpha-2-HS-glycoprotein and fetuin. FEBS Letts. 227, 167-170

Emori, Y., Kawasaki, H., Sugihara, H., Imajoh, S., Kawashima, S., & Suzuki, K. (1986) Isolation and sequence analyses of cDNA clones for the large subunits of two isozymes of rabbit calcium-dependent protease. J. Biol. Chem. 261, 9465-9471

Erikson, J., Martinis, J., & Croce, C.M. (1981) Assignment of the genes for human λ immunoglobulin chains to chromosome 22. Nature (London) 294, 173-175

Falquerho, L., Patey, G., Paquereau, L., Rossi, V., Lahuna, O., Szpirer, J., Szpirer, C., Levan, G., & Le Cam, A. (1991) Primary structure of the rat gene encoding an inhibitor of the insulin receptor tyrosine kinase. Gene 98, 209-216

Feinberg, A.P., & Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13

Fey, G.H., & Fuller, G.M. (1987) Regulation of acute phase gene expression by inflammatory mediators. Mol. Biol. Med. 4, 323-338

Fey, G.H., & Gauldie, J. (1990) The acute phase response of the liver in inflammation. In "Progress in liver disease, 9th edition". Eds. Popper, H., & Subramani, S., WB Saunders, Philadelphia.

Fink, D.J., & Gainer, H. (1980) Axonal transport of proteins. A new view using in vivo covalent labeling. J. Cell Biol. 85, 175-186

Fisher, D.A., & Lam, R.W. (1974) Thyroid hormone binding by bovine and ovine fetuin. Endocrinology 94, 49-54

Fisher, H.W., Puck, T.T., & Sato, G. (1958) Molecular growth requirements of single mammalian cells: the action of fetuin in promoting cell attachment to glass. Proc. Natl. Acad. Sci. USA 44, 4-10

Fisher, H.W., O'Brien, D., & Puck, T.T. (1962) The hydrolytic products of fetuin. Arch. Biochem. Biophys. 99, 241-248

Fletcher, C., Heintz, N., & Roeder, R.G. (1987) Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell 51, 773-781

Foucrier, J., Kraemer, M., Vassy, J., & Chalumeau, M.T. (1979) Demonstration of the simultaneous presence of transferrin, hemopexin and albumin in the same adult rat hepatocyte. Cell Differ. 8, 39-48

Fowlkes, D.M., Mullis, N.T., Comeau, C.M., & Crabtree, G.R. (1984) Potential basis for regulation of the coordinately expressed fibrinogen genes: homology in the 5' flanking regions. Proc. Natl. Acad. Sci. USA 81, 2313-2316

Frain, M., Swart, G., Monaci, P., Nicosia, A., Stämpfli, S., Frank, R., & Cortese, R. (1989) The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. Cell 59, 145-157

Fujita, T., Shibuya, H., Ohashi, T., Yamanishi, K., & Taniguchi, T. (1986) Regulation of human interleukin-2 gene: functional DNA sequences in the 5' flanking region for the gene expression in activated T lymphocytes. Cell 46, 401-407

Fukuma, Y., Kashimura, S., Umetsu, K., Yuasa, I., & Suzuki, T. (1990) Genetic variation of alpha-2-HS-glycoprotein in the Kyushu district of Japan: description of three new rare variants. Hum. Hered. 40, 49-51

Fukushima, Y., Bell, G.I., & Shows, T.B. (1988) The polymorphic human α_2-macroglobulin gene (A2M) is located in chromosome region 12p12.3→p13.3. Cytogenet. Cell Genet. 48, 58-59

Fuller, G.M., Bunzel, R.J., Woloski, B.M., & Nham, S.-U. (1987) Isolation of hepatocyte stimulating factor from human monocytes. Biochem. Biophys. Res. Commun. 144, 1003-1009

Fung, W.-P., & Schreiber, G. (1987) Structure and expression of the genes for major acute phase α_1-protein (thiostatin) and kininogen in the rat. J. Biol. Chem. 262, 9298-9308

Fung, W.-P., Thomas, T., Dickson, P.W., Aldred, A.R., Milland, J., Dziadek, M., Power, B., Hudson, P., & Schreiber, G. (1988) Structure and expression of the rat transthyretin (prealbumin) gene. J. Biol. Chem. 263, 480-488

Furuto-Kato, S., Matsumoto, A., Kitamura, N., & Nakanishi, S. (1985) Primary structure of the mRNAs encoding the rat precursors for bradykinin and T-kinin. Structural relationship of kininogens with major acute phase protein and α_1 cysteine proteinase inhibitor. J. Biol. Chem. 260, 12054-12059

Gaillard, D., Ailhaud, G., & Négrel, R. (1985) Fetuin modulates growth and differentiation of Ob17 preadipose cells in serum-free hormone-supplemented medium. Biochim. Biophys. Acta 846, 185-191

Galembeck, F., & Cann, J.R. (1974) Fetuin as a trypsin inhibitor. Arch. Biochem. Biophys. 164, 326-331

Gehring, M.R., Shiels, B.R., Northemann, W., de Bruijn, M.H.L., Kan, C.-C., Chain, A.C., Noonan, D.J., & Fey, G.H. (1987) Sequence of rat liver α_2-macroglobulin and acute phase control of its messenger RNA. J. Biol. Chem. 262, 446-454

Gejyo, F., & Schmid, K. (1981) Purification and characterization of the two forms of human plasma α_2HS-glycoprotein. Biochim. Biophys. Acta 671, 78-84

Gejyo, F., Chang, J., Bürgi, W., Schmid, K., Offner, G.D., Troxler, R., van Halbeek, H., Dorland, L., Gerwig, G.J., & Vliegenthart, J.F.G. (1983) Characterisation of the B-chain of human plasma α_2HS-glycoprotein. The complete amino acid sequence and primary structure of its heteroglycan. J. Biol. Chem. 258, 4966-4971

Gerard, G.F., & Miller, K. (1986) Comparison of glyoxal and formaldehyde gels for sizing rRNAs. BRL Focus 8:3, 5-6

Gerster, T., Matthias, P., Thali, M., Jiricny, J., & Schaffner, W. (1987) Cell type-specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J. 6, 1323-1330

Gilman, M.Z., Wilson, R.N., & Weinberg, R.A. (1986) Multiple protein-binding sites in the 5'-flanking region regulate c-*fos* expression. Mol. Cell. Biol. 6, 4305-4316

Gitschier, J., Drayna, D., Tuddenham, E.G.D., White, R.L., & Lawn, R.M. (1985) Genetic mapping and diagnosis of haemophilia A achieved through a *Bcl*I polymorphism in the factor VIII gene. Nature (London) 314, 738-740

Goldstein, I.M., Kaplan, H.B., Edelson, H.S., & Weissmann, G. (1982) Ceruloplasmin: an acute phase reactant that scavenges oxygen-derived free radicals. Ann. N.Y. Acad. Sci. 389, 368-379

Goldstein, I.M., Roos, D., Kaplan, H.B., & Weissmann, G. (1975) Complement and immunoglobulins stimulate superoxide production by human leukocytes independently of phagocytosis. J. Clin. Invest. 56, 1155-1163

Goodfellow, P.N., Jones, E.A., van Heyningen, V., Solomon, E., Bobrow, M., Miggiano, V., & Bodmer, W.F. (1975) The β_2 microglobulin gene is on chromosome 15 and not in the HL-A region. Nature (London) 254, 267-269

Gorin, M.B., Cooper, D.L., Eifermann, F., van de Rijn, P., & Tilghman, S.M. (1981) The evolution of α-fetoprotein and albumin. I. A comparison of the primary amino acid sequences of mammalian α-fetoprotein and albumin. J. Biol. Chem. 256, 1954-1959

Green, E.D., Adelt, G., Baenziger, J.U., Wilson, S., & van Halbeek, H. (1988) The asparagine-linked oligosaccharides on bovine fetuin. Strucutural analysis of *N*-glycanase-released oligosaccharides by 500-megahertz ^{1}H NMR spectroscopy. J. Biol. Chem. 263, 18253-18268

Green, W.A., & Kay, C.M. (1963) The influence of organic solvents and enzymatic modification on the secondary structure of fetuin. J. Biol. Chem. 238, 3640-3644

Greenberg, M.E., & Ziff, E.B. (1984) Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. Nature (London) 311, 433-438

Griffin, B.E., Soeda, E., Barrell, B.G., & Staden, R. (1980) Sequence and analysis of polyoma virus DNA. In "Molecular biology of tumor viruses, part II" Ed. Tooze, J., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 831-896

Grubb, A., & Löfberg, H. (1982) Human γ-trace, a basic microprotein: amino acid sequence and presence in the adenohypophysis. Proc. Natl. Acad. Sci. USA 79, 3024-3027

Grubb, A., & Löfberg, H. (1985) Human γ-trace. Structure, function and clinical use of concentration measurements. Scand. J. Clin. Lab. Invest. 45, Suppl. 177, 7-13

Grubb, A., Löfberg, H., & Barrett, A.J. (1984) The disulphide bridges of human cystatin C (γ-trace) and chicken cystatin. FEBS Letts. 170, 370-374

Haasemann, M., Nawratil, P., & Müller-Esterl, W. (1991) Rat tyrosine kinase inhibitor homologous to human α_2-HS glycoprotein and bovine fetuin. Biochem. J. 274, 899-902

Hajeri-Germond, M., Trojan, J., Uriel, J., & Hauw, J.J. (1984) In vitro uptake of exogenous α-fetoprotein by chicken dorsal root ganglia. Dev. Neurosci. 6, 111-117

Ham, R.G. (1964) Putrescine and related amines as growth factors for a mammalian cell line. Biochem. Biophys. Res. Commun. 14, 34-38

Hamberg, U., Elg, P., Nissinen, E., & Stelwagen, P. (1975) Purification and heterogeneity of human kininogen. Use of DEAE-chromatography, molecular sieving and antibody specific immunosorbents. Int. J. Peptide Protein Res. 7, 261-280

Hardon, E.M., Frain, M., Paonessa, G., & Cortese, R. (1988) Two distinct factors interact with the promoter regions of several liver-specific genes. EMBO J. 7, 1711-1719

Harper, M.E., & Dugaiczyk, A. (1983) Linkage of the evolutionarily-related serum albumin and α-fetoprotein genes within q11-22 of human chromosome 4. Am. J. Hum. Genet. 35, 565-572

Haymerle, H., Herz, J., Bressan, G.M., Frank, R., & Stanley, K.K. (1986) Efficient construction of cDNA libraries in plasmid expression vectors using an adaptor strategy. Nucleic Acids Res. 14, 8615-8624

von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133, 17-21

Heimburger, N., Heide, K., Haupt, H., & Schultze, H.E. (1964) Bausteinanalysen von Humanserumproteinen. Clin. Chim. Acta 10, 293-307

Heinrich, P.C., Castell, J.V., & Andus, T. (1990) Interleukin-6 and the acute phase reponse. Biochem. J. 265, 621-636

Henry, I., Uzan, G., Weil, D., Nicolas, H., Kaplan, J.C., Marguerie, C., Kahn, A., & Junien, C. (1984) The genes coding for the Aα-, Bβ-, and γ-chains of fibrinogen map to 4q2. Am. J. Hum. Genet. 36, 760-768

Heremans, J.F. (1960) Les globulines sérique du système gamma, pp. 103-107, Arscia, Brussels.

Herrera, R.E., Shaw, P.E., & Nordheim, A. (1989) Occupation of the c-*fos* serum response element *in vivo* by a multi protein complex is unaltered by growth factor induction. Nature (London) 340, 68-70

Hirado, M., Tsunasawa, S., Sakiyama, F., Niinobe, M., & Fujii, S. (1985) Complete amino acid sequence of bovine colostrum low-M_r cysteine proteinase inhibitor. FEBS Letts. 186, 41-45

Hobart, M.J., Rabbitts, T.H., Goodfellow, P.N., Solomon, E., Chambers, S., Spurr, N., & Povey, S. (1981) Immunoglobulin heavy chain genes in human are located on chromosome 14. Ann. Hum. Genet. 45, 331-335

Holmgren, R., Corces, V., Morimoto, R., Blackman, R., & Meselson, M. (1981) Sequence homologies in the 5' regions of four *Drosophila* heat-shock genes. Proc. Natl. Acad. Sci. USA 78, 3775-3778

Horwitz, J.P., Chua, J., Curby, R.J., Tomson, A.J., Da Rooge, M.A., Fisher, B.E., Mauricio, J., & Klundt, T. (1964) Substrates for cytochemical demonstration of enzyme activity. I. some substituted 3-indolyl-β-D-glycopyranosides. J. Med. Chem. 7, 574-575

Hsu, C.C.S., & Floyd, M. (1976) Lymphocyte stimulating and immunochemical properties of fetuin preparations. Protides Biol. Fluids 24, 295-302

Humphries, S.E., Williams, L., Myklebost, O., Stalenhoef, A.F.H., Demacker, P.N.M., Baggio, G., Crepaldi, G., Galton, D.J., & Williamson, R. (1984) Familial apolipoprotein CII deficiency: a preliminary analysis of the gene defect on two independent families. Hum. Genet. 67, 151-155

Imbra, R.J., & Karin, M. (1987) Metallothionein gene expression is regulated by serum factors and activators of protein kinase C. Mol. Cell. Biol. 7, 1358-1363

Ingolia, T.D., & Craig, E.A. (1981) Primary sequence of the 5' flanking regions of the Drosophila heat shock genes in chromosome subdivision 67B. Nucleic Acids Res. 9, 1627-1642

Ingolia, T.D., Craig, E.A., & McCarthy, B.J. (1980) Sequence of three copies of the gene for the major Drosophila heat shock induced protein and their flanking regions. Cell 21, 669-679

Isemura, S., Saitoh, E., & Sanada, K. (1986) Characterization of a new cysteine proteinase inhibitor of human saliva, cystatin SN, which is immunologically related to cystatin S. FEBS Letts. 198, 145-149

Ishiguro, H., Higashiyama, S., Ohkubo, I., & Sasaki, M. (1987) Mapping of functional domains of human high molecular weight and low molecular weight kininogens using murine monoclonal antibodies. Biochemistry 26, 7021-7029

Izban, M.G., & Papaconstantinou, J. (1989) Cell-specific expression of mouse albumin promoter. Evidence for cell-specific DNA elements within the proximal promoter region and cis-acting DNA elements upstream of -160. J. Biol. Chem. 264, 9171-9179

Jentoft, N. (1990) Why are proteins *O*-glycosylated ? Trends Biochem. Sci. 15, 291-294

Johnson, P.F., & McKnight, S.L. (1989) Eukaryotic transcriptional regulatory proteins. Ann. Rev. Biochem. 58, 799-839

Johnson, W.V., & Heath, E.C. (1986a) Evidence for posttranslational O-glycosylation of fetuin. Biochemistry 25, 5518-5525

Johnson, W.V., & Heath, E.C. (1986b) Structural features of bovine fetuin revealed from analysis of the primary translation product: anomalous behavior on sodium dodecyl sulfate-polyacrylamide gel electrophoresis is due largely to peptide and not solely to carbohydrate. Arch. Biochem. Biophys. 251, 732-737

Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J., & Tjian, R. (1987) A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell 48, 79-89

Jones, S.E. (1990) Developmental profile of a fetuin-like glycoprotein in neocortex, cerebrospinal fluid and plasma of postnatal tammar wallaby *(Macropus eugenii).* Ph.D. Thesis, University of Southampton, UK.

Jones, S.E., Dziegielewska, K.M., Saunders, N.R., Christie, D.L., Sueiras-Diaz, J., & Szelke, M. (1988) Early cortical plate specific glycoprotein in a marsupial species belongs to the same family as fetuin and α_2HS glycoprotein. FEBS Letts. 236, 411-414

Jones, S.E., Christie, D.L., Dziegielewska, K.M., Hinds, L.A., & Saunders, N.R. (1991) Developmental profile of a fetuin-like glycoprotein in neocortex, cerebrospinal fluid and plasma of post-natal tammar wallaby *(Macropus eugenii).* Anat. Embryol., 183, 313-320

Kampschmidt, R.F., Upchurch, H.F., Eddington, C.L., & Pulliam, L.A. (1973) Multiple biological activities of a partially purified leukocytic endogenous mediator. Am. J. Physiol. 224, 530-533

Kan, C.-C., Solomon, E., Belt, K.T., Chain, A.C., Hiorns, L.R., & Fey, G. (1985) Nucleotide sequence of cDNA encoding human α_2-macroglobulin and assignment of the chromosomal locus. Proc. Natl. Acad. Sci. USA 82, 2282-2286

Kanno, H., Huang, I.-Y., Kan, Y.W., & Yoshida, A. (1989) Two structural genes on different chromosomes are required for encoding the major subunit of human red cell glucose-6-phosphate dehydrogenase. Cell 58, 595-606

Kant, J.A., Fornace, A.J. Jr., Saxe, D., Simon, M.I., McBride, O.W., & Crabtree, G.R. (1985) Evolution and organization of the fibrinogen locus on chromosome 4: gene duplication accompanied by transposition and inversion. Proc. Natl. Acad. Sci. USA 82, 2344-2348

Karathanasis, S.K., McPherson, J., Zannis, V.I., & Breslow, J.L. (1983) Linkage of human apolipoproteins A-I and C-III genes. Nature (London) 304, 371-373

Karch, F., Török, I., & Tissières, A. (1981) Extensive regions of homology in front of two hsp70 heat shock variant genes in *Drosophila melanogaster.* J. Mol. Biol. 148, 219-230

Karin, M., & Richards, R.I. (1982) Human metallothionein genes - primary structure of the metallothionein-II gene and a related processed gene. Nature (London) 299, 797-802

Kassavetis, G.A., Butler, E.T., Roulland, D., & Chamberlin, M.J. (1982) Bacteriophage SP6-specific RNA polymerase. II. Mapping of SP6 DNA and selective *in vitro* transcription. J. Biol. Chem. 257, 5779-5788

Kato, A., Nakamura, Y., Miura, O., Hirosawa, S., Sumi, Y., & Aoki, N. (1988) Assignment of the human α_2-plasmin inhibitor gene (PLI) to chromosome region 18p11.1→q11.2 by in situ hybridization. Cytogenet. Cell Genet. 47, 209-211

Kaufman, K.M., Snider, J.V., Spurr, N.K., Schwartz, C.E., & Sodetz, J.M. (1989) Chromosomal assignment of genes encoding the α, β, and γ subunits of human complement protein C8: identification of a close physical linkage between the α and β loci. Genomics 5, 475-480

Kawasaki, E.S., & Wang, A.M. (1989) Detection of Gene Expression. In "PCR Technology: principles and applications for DNA amplfication" Ed. Erhlich, H.A., MacMillan Publishers Ltd., UK.

Kellermann, J., Haupt, H., Auerswald, E.-A., & Müller-Esterl, W. (1989) The arrangement of disulfide loops in human α_2-HS glycoprotein. Similarity to the disulfide bridge structures of cystatins and kininogens. J. Biol. Chem. 264, 14121-14128

Kellermann, J., Lottspeich, F., Henschen, A., & Müller-Esterl, W. (1986) Completion of the primary structure of human high-molecular-mass kininogen. The amino acid sequence of the entire heavy chain and evidence for its evolution by gene triplication. Eur. J. Biochem. 154, 471-478

Kellermann, J., Thelen, C., Lottspeich, F., Henschen, A., Vogel, R., & Müller-Esterl, W. (1987) Arrangement of the disulphide bridges in human low-M_r kininogen. Biochem. J. 247, 15-21

Kelly, K., Cochran, B.H., Stiles, C.D., & Leder, P. (1983) Cell-specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. Cell 35, 603-610

Kemp, B.E., & Pearson, R.B. (1990) Protein kinase recognition sequence motifs. Trends Biochem. Sci. 15, 342-346

Kioussis, D., Eiferman, F., van de Rijn, P., Gorin, M.B., Ingram, R.S., & Tilghman, S.M. (1981) The evolution of α-fetoprotein and albumin. II. The structures of the α-fetoprotein and albumin genes in the mouse. J. Biol. Chem. 256, 1960-1967

Kishi, K., Yasuda, T., & Ikehara, Y. (1988) Alpha-2-HS-glycoprotein polymorphism detected in human urine by isoelectric focusing and immunoblotting. Hum. Hered. 38, 207-210

Kitamura, N., Kitagawa, H., Fukushima, D., Takagaki, Y., Miyata, T., & Nakanishi, S. (1985) Structural organization of the human kininogen gene and a model for its evolution. J. Biol. Chem. 260, 8610-8617

Kitamura, N., Takagaki, Y., Furuto, S., Tanaka, T., Nawa, H., & Nakanishi, S. (1983) A single gene for bovine high molecular weight and low molecular weight kininogens. Nature (London) 305, 545-549

Kithier, K., Masopust, J., & Rádl, J. (1968) Fetal α-globulin of bovine serum differing from fetuin. Biochim. Biophys. Acta 160, 135-137

Koide, T. (1988) Human histidine-rich glycoprotein gene: evidence for evolutionary relatedness to cystatin supergene family. Thromb. Res. Suppl. VIII, 91-97

Koide, T., & Odani, S. (1987) Histidine-rich glycoprotein is evolutionarily related to the cystatin superfamily. Presence of two cystatin domains in the N-terminal region. FEBS Letts. 216, 17-21

Koide, T., Foster, D., Yoshitake, S., & Davie, E.W. (1986) Amino acid sequence of human histidine-rich glycoprotein derived from the nucleotide sequence of its cDNA. Biochemistry 25, 2220-2225

Kondo, H., Abe, K., Emori, Y., & Arai, S. (1991) Gene organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of the animal cystatins. FEBS Letts. 278, 87-90

Kondo, H., Emori, Y., Abe, K., Suzuki, K., & Arai, S. (1989) Cloning and sequence analysis of the genomic DNA fragment encoding oryzacystatin. Gene 81, 259-265

Kozak, M. (1989) The scanning model for translation: an update. J. Cell Biol. 108, 229-241

Kretsinger, R.H. (1987) Calcium coordination and the calmodulin fold: divergent versus convergent evolution. Cold Spring Harbor Symp. Quant. Biol. 52, 499-510

Krieg, P.A., & Melton, D.A. (1984) Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. Nucleic Acid Res. 12, 7057-7070

Krieg, P.A., & Melton, D.A. (1987) *In vitro* RNA synthesis with SP6 RNA polymerase. Methods Enzymol. 155, 397-415

Kruijer, W., Cooper, J.A., Hunter, T., & Verma, I.M. (1984) Platelet-derived growth factor induces rapid but transient expression of the c-*fos* gene. Nature (London) 312, 711-716

Kumbla, L., Cayatte, A.J., & Subbiah, M.T.R. (1989) Association of a lipoprotein-like particle with bovine fetuin. FASEB J. 3, 2075-2080

Kunz, D., Zimmermann, R., Heisig, M., & Heinrich, P.C. (1989) Identification of the promoter sequences involved in the interleukin-6 dependent expression of the rat α_2-macroglobulin gene. Nucleic Acids Res. 17, 1121-1138

Kurnit, D.M., Philipp, B.W., & Bruns, G.A.P. (1982) Confirmation of the mapping assignment of human serum albumin to chromosome 4 using a cloned human albumin gene. Cytogenet. Cell Genet. 34, 282-288

Kushner, I. (1982) The phenomenon of the acute phase response. Ann. N.Y. Acad. Sci. 389, 39-48

Kushner, I. (1988) The acute phase response: an overview. Methods Enzymol. 163, 373-383

Kuwano, R., Maeda, T., Usui, H., Araki, K., Yamakuni, T., Ohshima, Y., Kurihara, T., & Takahashi, Y. (1986) Molecular cloning of cDNA of S100α subunit mRNA. FEBS Letts. 202, 97-101

Laber, B., Krieglstein, K., Henschen A., Kos, J., Turk, V., Huber, R., & Bode, W. (1989) The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. FEBS Letts. 248, 162-168

Lachmann, P.J. (1982) Judging the state of the art. Immunol. Today 3, 27-28

Lackner, K.J., Law, S.W., Brewer, H.B. Jr., Sakaguchi, A.Y., & Naylor, S.L. (1984) The human apolipoprotein A-II gene is located on chromosome 1. Biochem. Biophys. Res. Commun. 122, 877-883

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227, 680-685

Lai, E.C., Kao, F.-T., Law, M.L., & Woo, S.L.C. (1983) Assignment of the α_1-antitrypsin gene and sequence-related gene to human chromosome 14 by molecular hybridization. Am. J. Hum. Genet. 35, 385-392

Landschulz, W.H., Johnson, P.F., Adashi, E.Y., Graves, B.J., & McKnight, S.L. (1988a) Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. 2, 786-800

Landschulz, W.H., Johnson, P.F., & McKnight, S.L. (1988b) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240, 1759-1764

Landschulz, W.H., Johnson, P.F., & McKnight, S.L. (1989) The DNA binding domain of the rat liver protein C/EBP is bipartite. Science 243, 1681-1688

Latchman, D.S. (1990) Eukaryotic transcription factors. Biochem J. 270, 281-289

Lau, L.F., & Nathans, D. (1985) Identification of a set of genes expressed during the GO/G1 transition of cultured mouse cells. EMBO J. 4, 3145-3151

Law, S.W., Lee, N., Monge, J.C., Brewer, H.B. Jr., Sakaguchi, A.Y., & Naylor, S.L. (1985) Human apoB-100 gene resides in the p23→pter region of chromosome 2. Biochem. Biophys. Res. Commun. 131, 1003-1012

Law, S.K.A., & Reid, K.B.M. (1988) Complement. IRL Press Ltd., Oxford, UK.

Lebreton, J.P. (1977) Purification of the human plasma alpha$_2$-SH glycoprotein by zinc chelate affinity chromatography. FEBS Letts. 80, 351-354

Lebreton, J.P., Joisel, F., Raoult, J.P., Lannuzel, B., Rogez, J.P., & Humbert, G. (1979) Serum concentration of human alpha$_2$ HS glycoprotein during the inflammatory process. Evidence that alpha$_2$ HS glycoprotein is a negative acute-phase reactant. J. Clin. Invest. 64, 1118-1129

Le Cam, A., & Le Cam, G. (1985) A new negatively regulated acute-phase phosphoprotein synthesized by rat hepatocytes. Biochem. J. 230, 603-607

Le Cam, A., Magnaldo, I., Le Cam, G., & Auberger, P. (1985) Secretion of a major phosphorylated glycoprotein by hepatocytes. Characterization of specific antibodies and investigations of the processing, excretion kinetics, and phosphorylation. J. Biol. Chem. 260, 15965-15971

Lee, C.-C. (1990) Human α_2-HS-glycoprotein: structure, function and tissue-specific gene expression. Ph.D. Thesis, University of Texas Graduate School of Biomedical Sciences at San Antonio, USA.

Lee, C.-C., Bowman, B.H., & Yang, F. (1987a) Human α_2-HS-glycoprotein: the A and B chains with a connecting sequence are encoded by a single mRNA transcript. Proc. Natl. Acad. Sci. USA 84, 4403-4407

Lee, W., Haslinger, A., Karin, M., & Tjian, R. (1987b) Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature (London) 325, 368-372

Lee, W., Mitchell, P., & Tjian, R. (1987c) Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell 49, 741-752

Lerman, M.I., & Colburn, N.H. (1987) Pro genes: a novel class of genes that specify sensitivity to induction of neoplastic transformation by tumor promoters. In "Tumor promotion: biological approaches for mechanistic studies and assay systems". Eds. Langenbach, R., Barrett, J.C., Elmore, E., Raven Press, New York.

Levin, M.J., Tuil, D., Uzan, G., Dreyfus, J.-C., & Kahn, A. (1984) Expression of the transferrin gene during development of non-hepatic tissue: high level of transferrin mRNA in fetal muscle and adult brain. Biochem. Biophys. Res. Commun. 122, 212-217

Lewis, J.G., & André, C.M. (1978) A serum DNA-binding protein absent in malignant disease. FEBS Letts. 92, 211-213

Lewis, J.G., & André, C.M. (1980) Effect of human α_2HS glycoprotein on mouse macrophage function. Immunol. 39, 317-322

Lewis, J.G., & André, C.M. (1981) Enhancement of human monocyte phagocytic function by α_2HS glycoprotein. Immunol. 42, 481-487

Lewis, J.G., & André, C.M. (1982) The α_2HS glycoprotein receptor on lymphocytes transformed by Epstein-Barr virus. FEBS Letts. 143, 332-336

Lewis, J.G., & André, C.M. (1984) The effect of α_2HS glycoprotein on lymphocyte reactivity to phytohemagglutinin. Immunol. Comm. 13, 35-47

Lewis, J.G., Crosier, P.S., & André, C.M. (1982) α_2HS glycoprotein binds to lymphocytes transformed by Epstein-Barr virus. FEBS Letts. 138, 37-39

Libby, P., Raines, E.W., Cullinane, P.M., & Ross, R. (1985) Analysis of the mitogenic effect of fetuin preparations on arterial smooth muscle cells: the role of contaminant platelet-derived growth factor. J. Cell. Physiol. 125, 357-366

Lichtsteiner, S., Wuarin, J. & Schibler, U. (1987) The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell 51, 963-973

Lim, H.M., & Pène, J.J. (1988) Optimal conditions for supercoil DNA sequencing with the *Escherichia coli* DNA polymerase I large fragment. Gene Anal. Techn. 5, 32-39

Lindquist, S. (1986) The heat-shock response. Ann. Rev. Biochem. 55, 1151-1191

Liu, J.-K., Bergman, Y., & Zaret, K.S. (1988) The mouse albumin promoter and a distal upstream site are simultaneously DNase I hypersensitive in liver chromatin and bind similar liver-abundant factors in vitro. Genes Dev. 2, 528-541

Lowe, T., Sharefkin, J., Yang, S.Q., & Dieffenbach, C.W. (1990) A computer program for selection of oligonucleotide primers for polymerase chain reactions. Nucleic Acids Res. 18, 1757-1761

Maeda, N., Yang, F., Barnett, D.R., Bowman, B.H., & Smithies, O. (1984) Duplication within the haptoglobin Hp^2 gene. Nature (London) 309, 131-135

Magnuson, V.L., McCombs, J.L., Lee, C.-C., Yang, F., Bowman, B.H., & McGill, J.R. (1988) Human α_2-HS-glycoprotein localized to 3q27→q29 by in situ hybridization. Cytogenet. Cell Genet. 47, 72-74

Maire, P., Wuarin, J., & Schibler, U. (1989) The role of cis-acting promoter elements in tissue-specific albumin gene expression. Science 244, 343-346

Malcolm, S., Barton, P., Murphy, C., Ferguson-Smith, M.A., Bentley, D.L., & Rabbitts, T.H. (1982) Localization of human immunoglobulin *k* light chain variable region genes to the short arm of chromosome 2 by *in situ* hybridization. Proc. Natl. Acad. Sci. USA 79, 4957-4961

Malone, J.D., Teitelbaum, S.L., Griffin, G.L., Senior, A.M., & Kahn, A.J. (1982) Recruitment of osteoclast precursors by purified bone matrix constituents. J. Cell Biol. 92, 227-230

Mancini, G., Carbonara, A.O., & Heremans, J.F. (1965) Immunochemical quantitation of antigens by single radial immunodiffusion. Immunochem. 2, 235-254

Mandel, M., & Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. 53, 159-162

Maniatis, T., Goodbourn, S., & Fischer, J.A. (1987) Regulation of inducible and tissue-specific gene expression. Science 236, 1237-1245

Mantzouranis, E.C., Dowton, S.B., Whitehead, A.S., Edge, M.D., Bruns, G.A.P., & Colten, H.R. (1985) Human serum amyloid P component: cDNA isolation, complete sequence of pre-serum amyloid P component, and localization of the gene to chromosome 1. J. Biol. Chem. 260, 7752-7756

Marinković, S., & Baumann, H. (1990) Structure, hormonal regulation, and identification of the interleukin-6- and dexamethasone-responsive element of the rat haptoglobin gene. Mol. Cell. Biol. 10, 1573-1583

Marr, A.G.M., Owen, J.A., & Wilson, G.S. (1962) Studies on the growth-promoting glycoprotein fraction of foetal calf serum. Biochim. Biophys. Acta 63, 276-285

Marti, J., Aliau, S., Bonfils, C., Vigne, C., & Moretti, J. (1973) Preparation, constantes physiques et constitution chimique de la fetuine d'agneau. Biochim. Biophys. Acta 303, 348-359

Matrisian, L.M., Leroy, P., Ruhlmann, C., Gesnel, M.-C., & Breathnach, R. (1986) Isolation of the oncogene and epidermal growth factor-induced transin gene: complex control in rat fibroblasts. Mol. Cell. Biol. 6, 1679-1686

Matsushima, K., Cheng, M., & Migita, S. (1982) Purification and physicochemical characterization of human α_2HS-glycoprotein. Biochim. Biophys. Acta 701, 200-205

Mbuyi, J.-M., Dequecker, J., Bloemmen, F., & Stevens, E. (1982) Plasma proteins in human cortical bone: enrichment of α_2 HS-glycoprotein, α_1 acid-glycoprotein, and IgE. Calcif. Tiss. Int. 34, 229-231

McBride, O.W., Battey, J., Hollis, G.F., Swan, D.C., Siebenlist, U., & Leder, P. (1982) Localization of human variable and constant region immunoglobulin heavy chain genes on subtelomeric band q32 of chromosome 14. Nucleic Acids Res. 10, 8155-8170

McCombs, J.L., Yang, F., Bowman, B.H., McGill, J.R., & Moore, C.M. (1986) Chromosomal localization of group-specific component by in situ hybridization. Cytogenet. Cell Genet. 42, 62-64

McGee, J.O'D., Rhoads, R.E., & Udenfriend, S. (1971) The substrate recognition site of collagen proline hydroxylase: the hydroxylation of -X-pro-gly- sequences in bradykinin analogs and other peptides. Arch. Biochem. Biophys. 144, 343-351

McGill, J.R., Yang, F., Baldwin, W.D., Brune, J.L., Barnett, D.R., Bowman, B.H., & Moore, C.M. (1984) Localization of the haptoglobin α and β genes *(HPA* and *HPB)* to human chromosome 16q22 by in situ hybridization. Cytogenet. Cell Genet. 38, 155-157

Miller, H. (1987) Practical aspects of preparing phage and plasmid DNA: growth, maintenance, and storage of bacteria and bacteriophage. Methods Enzymol. 152, 145-170

Miller, L.L., Bly, C.G., Watson, M.L., & Bale, W.F. (1951) The dominant role of the liver in plasma protein synthesis. A directed study of the isolated perfused rat liver with the aid of lysine-ϵ-C^{14}. J. Exp. Med. 9, 431-453

Mohun, T.J., Garrett, N., & Treisman, R.H. (1987) *Xenopus* cytoskeletal actin and human c-*fos* gene promoters share a conserved protein-binding site. EMBO J. 6, 667-673

Møllgård, K, & Saunders, N.R. (1986) The development of the human blood-brain and blood-CSF barrier. Neuropath. & Appl. Neurobiol. 12, 337-358

Møllgård, K., Dziegielewska, K.M., Saunders, N.R., Zakut, H., & Soreq, H. (1988) Synthesis and localization of plasma proteins within cell types in the developing human brain. Integrity of the fetal blood-brain barrier to endogenous proteins of hepatic origin. Dev. Biol. 128, 207-221

Møllgård, K., Reynolds, M.L., Jacobsen, M., Dziegielewska, K.M., & Saunders, N.R. (1984) Differential immunocytochemical staining for fetuin and transferrin in the developing cortical plate. J. Neurocytol. 13, 497-502

Monaci, P., Nicosia, A., & Cortese, R. (1988) Two different liver-specific factors stimulate *in vitro* transcription from the human α1-antitrypsin promoter. EMBO J. 7, 2075-2087

Moore, C.M., McCombs, J.L., & McGill, J.R. (1985) *In situ* hybridization for chromosomal mapping of plasma proteins. Protides Biol. Fluids 35, 83-86

Morgan, W.D., Williams, G.T., Morimoto, R.I., Greene, J., Kingston, R.E., & Tjian, R. (1987) Two transcriptional activators, CCAAT-box binding transcription factor and heat shock transcription factor, interact with a human hsp70 gene promoter. Mol. Cell. Biol. 7, 1129-1138

Moro, R., Fielitz, W., Esteves, A., Grunberg, J., & Uriel, J. (1984) *In vivo* uptake of heterologous alphafetoprotein and serum albumin by ependymal cells of developing chick embryos. Int. J. Dev. Neurosci. 2, 143-148

Motojima, K., & Goto, S. (1989) Brain-specific expression of transthyretin mRNA as revealed by cDNA cloning from brain. FEBS Letts. 258, 103-105

Motojima, K., & Goto, S. (1990) Dual promoters and tissue-specific expression of rat transthyretin gene. Biochem. Biophys. Res. Commun. 173, 323-330

Müller, R., Bravo, R., Burckhardt, J., & Curran, T. (1984) Induction of c-*fos* gene and protein by growth factors precedes activation of c-*myc*. Nature (London) 312, 716-720

Müller-Eberhard, H.J. (1975) The complement system. In "The Plasma Proteins. Vol. 1. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Müller-Esterl, W., Fritz, H., Kellermann, J., Lottspeich, F., Machleidt, W., & Turk, V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. FEBS Letts. 191, 221-226

Murray, A.C., Oikawa, K., & Kay, C.M. (1969) Circular dichroism studies on native fetuin and some of its derivatives. Biochim. Biophys. Acta 175, 331-338

Murray, J.C., Watanabe, K., Tamaoki, T., Hornung, S., & Motulsky, A. (1985) RFLP's for human alphafetoprotein *(AFP)* at 4q11-4q13. Nucleic Acids Res. 13, 6794

Nakai, A., Hirayoshi, K., Saga, S., Yamada, K.M., & Nagata, K. (1989) The transformation-sensitive heat shock protein (HSP47) binds specifically to fetuin. Biochem. Biophys. Res. Commun. 164, 259-264

Nakanishi, S., Kitamura, N., Ohkubo, H., Kakizuka, A., Kageyama, R., Masu, Y., & Nakayama, K. (1987) Control of kininogen gene expression. Adv. Exp. Med. Biol. 247A, 7-18

Nawa, H., Kitamura, N., Hirose, T., Asai, M., Inayama, S., & Nakanishi, S. (1983) Primary structures of bovine liver low molecular weight kininogen precursors and their two mRNAs. Proc. Natl. Acad. Sci. USA 80, 90-94

Nilsson, B., Nordén, N.E., & Svensson, S. (1979) Structural studies on the carbohydrate portion of fetuin. J. Biol. Chem. 254, 4545-4553

Northoff, H., Andus, T., Tran-Thi, T.-A., Bauer, J., Decker, K., Kubanek, B., & Heinrich, P.C. (1987) The inflammation mediators interleukin 1 and hepatocyte-stimulating factor are differently regulated in human monocytes. Eur. J. Immunol. 17, 707-711

Oberholtzer, J.C., Englander, S.W., & Horwitz, A.F. (1981) Hydrogen-bonded structure of the complex N-linked fetuin glycopeptide. Biochemistry 20, 4785-4792

Ohkubo, I., Kurachi, K., Takasawa, T., Shiokawa, H., & Sasaki, M. (1984) Isolation of a human cDNA for α_2-thiol proteinase inhibitor and its identity with low molecular weight kininogen. Biochemistry 23, 5691-5697

Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., & Matsubara, K. (1985) An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. J. Biol. Chem. 260, 2605-2608

Olaisen, B., Gedde-Dahl, T. Jr., Teisberg, P., Thorsby, E., Siverts, A., Jonassen, R., & Wilhelmy, M.C. (1985) A structural locus for coagulation factor XIIIA *(F13A)* is located distal to the HLA region on chromosome 6p in man. Am. J. Hum. Genet. 37, 215-220

Olaisen, B., Teisberg, P., & Gedde-Dahl, T. Jr. (1982) The locus for apolipoprotein E (apoE) is linked to the complement component C3 (C3) locus on chromosome 19 in man. Hum. Genet. 62, 233-236

Olaisen, B., Teisberg, P., Jonassen, P., Thorsby, E., & Gedde-Dahl, T. Jr. (1983) Gene order and gene distances in the HLA region studied by the haplotype method. Ann. Hum. Genet. 47, 285-292

Oliviero, S., & Cortese, R. (1989) The human haptoglobin gene promoter: interleukin-6-responsive elements interact with a DNA-binding protein induced by interleukin-6. EMBO J. 8, 1145-1151

Oliviero, S., Morrone, G., & Cortese, R. (1987) The human haptoglobin gene: transcriptional regulation during development and acute phase induction. EMBO J. 6, 1905-1912

Orlandi, R., Güssow, D.H., Jones, P.T., & Winter, G. (1989) Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA 86, 3833-3837

Oshiro, Y., & Eylar, E.H. (1968) Physical and chemical studies on glycoproteins. III. The microheterogeneity of fetuin, a fetal calf serum glycoprotein. Arch. Biochem. Biophys. 127, 476-489

Oshiro, Y., & Eylar, E.H. (1969) Physical and chemical studies on glycoproteins. IV. The influence of sialic acid on the conformation of fetuin. Arch. Biochem. Biophys. 130, 227-234

van Oss, C.J., & Bronson, P.M. (1974) Isolation of human serum alpha-2 HS glycoprotein. Prep. Biochem. 4, 127-139

van Oss, C.J., Bronson, P.M., & Border, J. R. (1975) Changes in the serum alpha glycoprotein distribution in trauma patients. J. Trauma 15, 451-455

van Oss, C.J., Gillman, C.F., Bronson, P.M., & Border, J.R. (1974) Opsonic properties of human alpha-2 HS glycoprotein. Immunol. Comm. 3, 329-335

Pearson, S., Tetri, P., & Francke, U. (1982) Chromosome 14 codes for human α_1-antitrypsin (PI) expression in rat hepatoma x human fetal liver cell hybrids. Cytogenet. Cell Genet. 32, 309

Pedersen, K.O. (1944) Fetuin, a new globulin isolated from serum. Nature (London) 154, 575

Pelham, H.R.B. (1982) A regulatory upstream promoter element in the Drosophila Hsp 70 heat-shock gene. Cell 30, 517-528

Pennica, D., Holmes, W.E., Kohr, W.J., Harkins, R.N., Vehar, G.A., Ward, C.A., Bennett, W.F., Yelverton, E., Seeburg, P., Heyneker, H.L., Goeddel, D.V., & Collen, D. (1983) Cloning and expression of human tissue-type plasminogen activator cDNA in *E. coli*. Nature (London) 301, 214-221

Pepys, M.B., & Baltz, M.L. (1983) Acute phase proteins with special reference to C-reactive protein and related proteins (pentaxins) and serum amyloid A protein. Adv. Immunol. 34, 141-212

Pierce, M., Cummings, R.D., Cebula, T.A., & Roth, S. (1979) Galactosyltransferase from commercial preparations of fetuin. Biochim. Biophys. Acta 571, 166-170

Poli, V., & Cortese, R. (1989) Interleukin 6 induces a liver-specific nuclear protein that binds to the promoter of acute-phase genes. Proc. Natl. Acad. Sci. USA 86, 8202-8206

Poli, V., Mancini, F.P., & Cortese, R. (1990) IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP. Cell 63, 643-653

Poupart, P., Vandenabeele, P., Cayphas, S., van Snick, J., Haegeman, G., Kruys, V., Fiers, W., & Content, J. (1987) B cell growth modulating and differentiating activity of recombinant human 26-kd protein (BSF-2, HuIFN-β_2, HPGF). EMBO J. 6, 1219-1224

Powanda, M.C., Cockerell, G.L., & Pekarek, R.S. (1973) Amino acid and zinc movement in relation to protein synthesis early in inflammation. Am. J. Physiol. 225, 399-401

Powell, D.J., Friedman, J.M., Oulette, A.J., Krauter, K.S., & Darnell, J.E. Jr. (1984) Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. J. Mol. Biol. 179, 21-35

Proudfoot, N.J., & Brownlee, G.G. (1976) 3' non-coding region sequences in eukaryotic messenger RNA. Nature (London) 263, 211-214

Prowse, K.R., & Baumann, H. (1988) Hepatocyte-stimulating factor, β_2 interferon, and interleukin-1 enhance expression of the rat α_1-acid glycoprotein gene via a distal upstream regulatory region. Mol. Cell. Biol. 8, 42-51

Prywes, R., & Roeder, R.G. (1986) Inducible binding of a factor to the c-*fos* enhancer. Cell 47, 777-784

Ptashne, M. (1988) How eukaryotic transcriptional activators work. Nature (London) 335, 683-689

Puck, T.T., Waldren, C.A., & Jones, C. (1968) Mammalian cell growth proteins I. growth stimulation by fetuin. Proc. Natl. Acad. Sci. USA 59, 192-199

Putkey, J.A., Ts'ui K.F., Tanaka, T., Lagacé, L., Stein, J.P., Lai, E.C., & Means, A.R. (1983) Chicken calmodulin genes. A comparison of cDNA sequences and isolation of a genomic clone. J. Biol. Chem. 258, 11864-11870

Putnam, F.W. (1987) Immunoglobulins: structure, function, and genes. In "The Plasma Proteins. Vol. 5. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Quelch, K.J., Cole, W.G., & Melick, R.A. (1984) Noncollagenous proteins in normal and pathological human bone. Calcif. Tiss. Int. 36, 545-549

Rawlings, N.D., & Barrett, A.J. (1990) Evolution of proteins of the cystatin superfamily. J. Mol. Evol. 30, 60-71

Raymondjean, M., Cereghini, S., & Yaniv, M. (1988) Several distinct "CCAAT" box binding proteins coexist in eukaryotic cells. Proc. Natl. Acad. Sci. USA 85, 757-761

Reid, K.B.M., & Porter, R.R. (1981) The proteolytic activation systems of complement. Ann. Rev. Biochem. 50, 433-464

Reynolds, M.L., Sarantis, M.E.P., Lorscheider, F.L., & Saunders, N.R. (1987) Fetuin as a marker of cortical plate cells in the fetal cow neocortex: a comparison of the distribution of fetuin, α_2HS-glycoprotein, α-fetoprotein and albumin during early development. Anat. Embryol. 175, 355-363

Rizzino, A., & Sato, G. (1978) Growth of embryonal carcinoma cells in serum-free medium. Proc. Natl. Acad. Sci. USA 75, 1844-1848

Rocchi, M., Roncuzzi, L., Santamaria, R., Sbarra, D., Mochi, M., Archidiacono, N., Covone, A., Cortese, R., & Romeo, G. (1986) Mapping of coagulation factors protein C and factor X on chromosome 2 and 13 respectively. Cytogenet. Cell Genet. 40, 734

Rogde, S., Olaisen, B., Gedde-Dahl, T. Jr., & Teisberg, P. (1986) The C8A and C8B loci are closely linked on chromosome 1. Ann. Hum. Genet. 50, 139-144

Rogne, S., Myklebost, O., Hoyheim, B., Olaisen, B., & Gedde-Dahl, T. Jr. (1989a) The genes for apolipoprotein AII (APOA2) and the Duffy blood group (FY) are linked on chromosome 1 in man. Genomics 4, 169-173

Rogne, S., Myklebost, O., Stanley, K., & van Kessel, A.G. (1989b) The gene for human complement C9 is on chromosome 5. Genomics 5, 149-152

Rohrlich, S.T., & Rifkin, D.B. (1981) Isolation of the major serine protease inhibitor from the 5-day serum-free conditioned medium of human embryonic lung cells and demonstration that it is fetuin. J. Cell. Physiol. 109, 1-15

Rosales, R., Vigneron, M., Macchi, M., Davidson, I., Xiao, J.H., & Chambon, P. (1987) *In vitro* binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. EMBO J. 6, 3015-3025

Saitoh, E., Isemura, S., Sanada, K., Kim, H.-S., Smithies, O. & Maeda, N. (1988) Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily related to family III cystatin genes. Biol. Chem. Hoppe-Seyler 369, (suppl.) 191-197

Saitoh, E., Sabatini, L.M., Eddy, R.L., Shows, T.B., Azen, E.A., Isemura, S., & Sanada, K. (1989) The human cystatin C gene (CST3) is a member of the cystatin gene family which is localized on chromosome 20. Biochem. Biophy. Res. Commun. 162, 1324-1331

Sakane, F., Yamada, K., Kanoh, H., Yokoyama, C., & Tanabe, T. (1990) Porcine diacylglycerol kinase sequence has zinc finger and E-F hand motifs. Nature (London) 344, 345-348

Salomon, D.S., Bano, M., Smith, K.B., & Kidwell, W.R. (1982) Isolation and characterization of a growth factor (embryonin) from bovine fetuin which resembles α_2-macroglobulin. J. Biol. Chem. 257, 14093-14101

Salomon, D.S., Smith, K.B., Losonczy, I., Bano, M., Kidwell, W.R., Alessandri, J., & Gullino, P.M. (1984) α_2-macroglobulin, a contaminant of commercially prepared Pedersen fetuin: isolation, characterization, and biological activity. In "Cell culture methods for molecular and cell biology" Eds. Barnes, D.W., Sirbasku, D.A., & Sato, G.H., Alan R. Liss, Inc., New York., Vol. 3, pp125-153

Salvesen, G., Parkes, C., Abrahamson, M., Grubb, A., & Barrett, A.J. (1986) Human low-M_r kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. Biochem. J. 234, 429-434

Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) Molecular cloning: a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467

Sarantis, M.E.P., & Saunders, N.R. (1986) An immunocytochemical demonstration of α_2HS-glycoprotein in the developing neocortex of the rat. Neurosci. Letts. 65, 346-350

Scanu, A.M., Edelstein, C., & Keim, P. (1975) Serum lipoproteins. In "The Plasma Proteins. Vol. 1. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Scanu, A.M. (1987) Plasma apolipoproteins: gene structure, function, and variants. In "The Plasma Proteins. Vol. 5. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Schachter, B.S., & Toran-Allerand, C.D. (1982) Intraneuronal α-fetoprotein and albumin are not synthesised locally in the developing brain. Dev. Brain Res. 5, 93-98

Schamaun, O., Olaisen, B., Mevåg, B., Gedde-Dahl, T. Jr., Ehnholm, C., & Teisberg, P. (1984) The two apolipoprotein loci *apoA-I* and *apoA-IV* are closely linked in man. Hum. Genet. 68, 181-184

Schelp, F.P., Thanangkul, O., Supawan, V., & Pongpaew, P. (1980) α_2HS-glycoprotein serum levels in protein-energy malnutrition. Br. J. Nutr. 43, 381-383

Schmid, K., & Bürgi, W. (1961) Preparation and properties of the human Ba-α_2-glycoproteins. Biochim. Biophys. Acta 47, 440-453

Schorpp, M., Kugler, W., Wagner, U., & Ryffel, G.U. (1988) Hepatocyte-specific promoter element HP1 of the *Xenopus* albumin gene interacts with transcriptional factors of mammalian hepatocytes. J. Mol. Biol. 202, 307-320

Schreiber, G. (1987) Synthesis, processing, and secretion of plasma proteins by the liver and other organs and their regulation. In "The Plasma Proteins. Vol. 5. Structure, function and genetic control" Ed. Putnam, F.W., Academic Press Inc., London.

Schreiber, G., Tsykin, A., Aldred, A.R., Thomas, T., Fung, W.-P., Dickson, P.W., Cole, T., Birch, H., de Jong, F.A., & Milland, J. (1989) The acute phase response in the rodent. Ann. N.Y. Acad. Sci. 557, 61-86

Schultze, H.E., Heide, K., & Haupt, H. (1962) Charakterisierung eines niedermolekularen α_2-Mukoids aus Humanserum. Naturwiss. 49, 15

Schwabe, C., Anastasi, A., Crow, H., McDonald, J.K., & Barrett, A.J. (1984) Cystatin. Amino acid sequence and possible secondary structure. Biochem J. 217, 813-817

Sebetan, I.M. (1988) A simple pattern method for alpha$_2$ HS-glycoprotein typing. J. Forensic Sci. 33, 1031-1034

Sebetan, I.M., & Heshmat, M.M. (1988) Genetic polymorphism of desialyzed alpha$_2$ HS-glycoprotein by ultrathin isoelectric focusing. Z. Rechtsmed. 101, 205-207

Sehgal, P.B., May, L.T., Tamm, I., & Vilček, J. (1987) Human β_2 interferon and B-cell differentation factor BSF-2 are identical. Science 235, 731-732

Sell, S., Longley, M.A., & Boulter, J. (1985) α-fetoprotein and albumin gene expression in brain and other tissues of fetal and adult rats. Dev. Brain Res. 22, 49-53

Shiang, R., Murray, J.C., Morton, C.C., Buetow, K.H., Wasmuth, J.J., Olney, A.H., Sanger, W.G., & Goldberger, G. (1989) Mapping of the human complement factor I gene to 4q25. Genomics 4, 82-86

de Simone, V., Ciliberto, G., Hardon, E., Paonessa, G., Palla, F., Lundberg, L., & Cortese, R. (1987) *Cis*- and *trans*-acting elements responsible for the cell-specific expression of the human α1-antitrypsin. EMBO J. 6, 2759-2766

Slaga, T.J. (1983) Cellular and molecular mechanisms of tumour promotion. Cancer Surveys 2, 595-612

Smith, A.J., Matthews, J.B., Wilson, C., & Sewell, H.F. (1985) Plasma proteins in human cortical bone: *in vitro* binding studies. Calcif. Tiss. Int. 37, 208-210

Smith, E.L., Hill, R.L., Lehman, I.R., Lefkowitz, R.J., Handler, P., & White, A. (1983) Principles of Biochemistry. Mammalian Biochemistry. Seventh edition. McGraw-Hill Book Co., Singapore.

van Snick, J. (1990) Interleukin-6: an overview. Ann. Rev. Immunol. 8, 253-278

Solomon, E., Skok, J., Griffin, J., & Reid, K.B.M. (1986) Human C1q B chain (C1QB) is on chromosome 1p. Cytogenet. Cell Genet. 40, 749

Sorger, P.K., & Pelham, H.R.B. (1987) Purification and characterization of a heat-shock element binding protein from yeast. EMBO J. 6, 3035-3041

Sorger, P.K., & Pelham, H.R.B. (1988) Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. Cell 54, 855-864

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517

Sparrow, J.R. (1981) Immunocytochemical localization of plasma proteins in neuronal perikarya. Brain Res. 212, 159-163

Spiro, M.J., & Spiro, R.G. (1962) Composition of the peptide portion of fetuin. J. Biol. Chem. 237, 1507-1510

Spiro, R.G. (1960) Studies on fetuin, a glycoprotein of fetal serum. I. isolation, chemical composition, and physicochemical properties. J. Biol. Chem. 235, 2860-2869

Spiro, R.G. (1963) Demonstration of a single peptide chain in the glycoprotein fetuin: terminal amino acid analyses and studies of the oxidized and reduced alkylated preparations. J. Biol. Chem. 238, 644-649

Spiro, R.G., & Bhoyroo, V.D. (1974) Structure of the *O*-glycosidically linked carbohydrate units of fetuin. J. Biol. Chem. 249, 5704-5717

Splitter, G.A., & Everlith, K.M. (1982) Suppression of bovine T- and B-lymphocyte responses by fetuin, a bovine glycoprotein. Cell. Immunol. 70, 205-218

Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarčič, B., & Turk, V. (1990) The refined 2.4Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J. 9, 1939-1947

Sturm, R., & Herr, W. (1988) The POU domain is a bipartite DNA-binding structure. Nature (London) 336, 601-604

Subramaniam, M., Schmidt, L.J., Crutchfield, C.E. III, & Getz, M.J. (1989) Negative regulation of serum-responsive enhancer elements. Nature (London) 340, 64-66

Sueyoshi, T., Miyata, T., Kato, H., & Iwanaga, S. (1984) Disulfide bonds in bovine HMW kininogens. Seikagaku 56, 808-??

Sueyoshi, T., Enjyoji, K.-I., Shimada, T., Kato, H., Iwanaga, S., Bando, Y., Kominami, E., & Katunuma, N. (1985) A new function of kininogens as thiol-proteinase inhibitors: inhibition of papain and cathepsins B, H and L by bovine, rat and human plasma kininogens. FEBS Letts. 182, 193-195

Syvänen, A.-C., Bengtström, M., Tenhunen, J., & Söderlund, H. (1988) Quantification of polymerase chain reaction products by affinity-based hybrid collection. Nucleic Acids Res. 16, 11327-11338

Tabor, S., & Richardson, C.C. (1989) Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. J. Biol. Chem. 264, 6447-6458

Takasaki, S., & Kobata, A. (1986) Asparagine-linked sugar chains of fetuin: occurrence of tetrasialyl trinantennary sugar chains containing the Galβ1→3GlcNAc sequence. Biochemistry 25, 5709-5715

Takio, K., Kominami, E., Bando, Y., Katunuma, N., & Titani, K. (1984) Amino acid sequence of rat epidermal thiol proteinase inhibitor. Biochem. Biophys. Res. Commun. 121, 149-154

Takio, K., Kominami, E., Wakamatsu, N., Katanuma, N., & Titani, K. (1983) Amino acid sequence of rat liver thiol proteinase inhibitor. Biochem. Biophys. Res. Commun. 115, 902-908

Tata, F., Henry, I., Markham, A.F., Wallis, S.C., Weil, D., Grzeschik, K.H., Junien, C., Williamson, R., & Humphries, S.E. (1985) Isolation and characterisation of a cDNA clone for human apolipoprotein CI and assignment of the gene to chromosome 19. Hum. Genet. 69, 345-349

Thomas, A.S. (1989a) Phenotyping of α-2-HS glycoprotein in blood-stains by isoelectric focusing and immunoblotting. J. Forensic Sci. Soc. 29, 325-330

Thomas, A.S. (1989b) The evaluation of five electrophoretic phenotyping systems for routine screening of bloodstains. J. Forensic Sci. Soc. 29, 243-248

Thomas, G.H., & Elgin, S.C.R. (1988) Protein/DNA architecture of the DNase I hypersensitive region of the *Drosophila hsp26* promoter. EMBO J. 7, 2191-2201

Thomas, T., Schreiber, G., & Jaworowski, A. (1989) Developmental patterns of gene expression of secreted proteins in brain and choroid plexus. Dev. Biol. 134, 38-47

Thorbecke, G.J., Huriemann, J., & Silverstein, A.M. (1967) Production of fetuin and other serum proteins by fetal sheep liver *in vitro*. Proc. Soc. Exptl. Biol. 126, 816-819

Török, I., & Karch, F. (1980) Nucleotide sequences of heat shock activated genes in Drosophila melanogaster. I. Sequences in the regions of the 5' and 3' ends of the hsp 70 gene in the hybrid plasmid 56H8. Nucleic Acids Res. 8, 3105-3123

Townsend, R.R., Hardy, M.R., Wong, T.C., & Lee, Y.C. (1986) Binding of N-linked bovine fetuin glycopeptides to isolated rabbit hepatocytes: Gal/GalNAc hepatic lectin discrimination between Galβ(1,4)GlcNAc and Galβ(1,3)GlcNAc in a triantennary structure. Biochemistry 25, 5716-5725

Treisman, R. (1985) Transient accumulation of c-*fos* RNA following serum stimulation requires a conserved 5' element and c-*fos* 3' sequences. Cell 42, 889-902

Treisman, R. (1986) Identification of a protein-binding site that mediates transcriptional response of the c-*fos* gene to serum factors. Cell 46, 567-574

Treisman, R. (1987) Identification and purification of a polypeptide that binds to the c-*fos* serum response element. EMBO J. 6, 2711-2717

Triffitt, J.T., Gebauer, U., Ashton, B.A., Owen, M.E., & Reynolds, J.J. (1976) Origin of plasma α_2HS-glycoprotein and its accumulation in bone. Nature (London) 262, 226-227

Triffitt, J.T., & Owen, M. (1973) Studies on bone matrix glycoproteins. Incorporation of [1-^{14}C]glucosamine and plasma [^{14}C]glycoprotein into rabbit cortical bone. Biochem. J. 136, 125-134

Tsuchiya, Y., Hattori, M., Hayashida, K., Ishibashi, H., Okubo, H., & Sakaki, Y. (1987) Sequence analysis of the putative regulatory region of the rat α_2-macroglobulin gene. Gene 57, 73-80

Tufty, R.M., & Kretsinger, R.H. (1975) Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. Science 187, 167-169

Turk, V., Brzin, J., Kotnik, M., Lenarčič, B., Popović, T., Ritonja, A., Trstenjak, M., Begić-Odorbašić, L., & Machleidt, W. (1986) Human cysteine proteinases and their protein inhibitors: stefins, cystatins and kininogens. Biomed. Biochim. Acta 45, 1375-1384

Turk, V., Brzin, J., Longer, M., Ritonja, A., Eropkin, M., Borchart, U., & Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496

Umetsu, K., Kashimura, S., Ikeda, N., & Suzuki, T. (1984a) A new α_2HS-glycoprotein allele *(AHS*5)* in two Japanese families. Hum. Genet. 68, 264-265

Umetsu, K., Kashimura, S., Ikeda, N., & Suzuki, T. (1984b) A new α_2HS-glycoprotein typing by isoelectric focusing. Hum. Genet. 67, 70-71

Umetsu, K., Yuasa, Y., & Suzuki, T. (1986) The polymorphism of desialyzed α_2HS-glycoprotein (AHS): isoelectric focusing in 2.5*M* urea as a method for identification of genetic variants. Hum. Genet. 73, 372-373

Umetsu, K., Yuasa, I., Nishimura, H., Sasaki, H., & Suzuki, T. (1988) Genetic polymorphisms of orosomucoid and alpha-2-HS-glycoprotein in a Philippine population. Hum. Hered. 38, 287-290

Uriel, J., Faivre-Bauman, A., Trojan, J., & Foiret, D. (1981) Immunocytochemical demonstration of alpha-fetoprotein uptake by primary cultures of fetal hemisphere cells from mouse brain. Neurosci. Letts. 27, 171-175

Verpoorte, J.A., & Kay, C.M. (1966) Optical rotatory dispersion and spectrophotometric studies on fetuin in solutions containing detergent and denaturing reagents. Biochim. Biophys. Acta 126, 551-569

Verpoorte, J.A., Green, W.A., & Kay, C.M. (1963) Molecular weight, optical rotation, and tryptic susceptibility studies on modified fetuins. J. Biol. Chem. 240, 1156-1161

Wallace, M.R., Naylor, S.L., Kluve-Beckerman, B., Long, G., McDonald, L., Shows, T.B., & Benson, M. (1985) Localization of the human prealbumin gene to chromosome 18. Biochem. Biophys. Res. Commun. 129, 753-758

Wang, H., Koschinsky, M., & Hamerton, J.L. (1988) Localization of the processed gene for human ceruloplasmin to chromosome region 8q21.13→q23.1 by in situ hybridization. Cytogenet. Cell Genet. 47, 230-231

Wannemacher, R.W. Jr., Pekarek, R.S., Thompson, W.L., Curnow, R.T., Beall, F.A., Zenser, T.V., DeRubertis, F.R., & Beisel, W.R. (1975) A protein from polymorphonuclear leukocytes (LEM) which affects the rate of hepatic amino acid transport and synthesis of acute-phase globulins. Endocrin. 96, 651-661

Wasylyk, B. (1988) Enhancers and transcription factors in the control of gene expression. Biochim. Biophys. Acta 951, 17-35

Webb, G.C., Earle, M.E., Merritt, C., & Board, P.G. (1988) Localization of human α_1 acid glycoprotein genes to 9q31→q34.1 Cytogenet. Cell Genet. 47, 18-21

Weidinger, S., Schwarzfischer, F., Burgemeister, R., & Cleve, H. (1984) Genetic alpha-2-HS glycoprotein phenotypes demonstrated by isoelectric focusing and immunofixation. In: "Electrophoresis '84" Ed. Neuhoff, V., Verlag Chemie, Weinheim, Germany. pp. 487-490.

Weinstein, I.B., Lee, L.-S., Fisher, P.B., Mufson, A., & Yamasaki, H. (1979) Action of phorbol esters in cell culture: mimicry of transformation, altered differentiation, and effects on cell membranes. J. Supramol. Struct. 12, 195-208

Weitkamp, L.R. (1983) Evidence for linkage between the loci for transferrin and ceruloplasmin in man. Ann. Hum. Genet. 47, 293-297

Weitkamp, L.R., Renwick, J.H., Berger, J., Shreffler, D.C., Drachmann, O., Wuhrmann, R., Braend, M., Franglen, G. (1970) Additional data and summary for albumin-Gc linkage in man. Hum. Hered. 20, 1-7

Westwood, S.A. (1988) The use of α_2HS-glycoprotein and group specific component in typing forensic blood samples for discriminative and investigative purposes. Electrophoresis 9, 432-437

Westwood, S.A., Seaman, P.J., Ablett, P.J., Yuasa, I., Weidinger S., & Umetsu, K. (1987a) *A2HS*11:* a new allele of alpha-2-HS-glycoprotein found in Afro-Caribbeans. Electrophoresis 8, 559-561

Westwood, S.A., Seaman, P.J., O'Brien, C., & Thorogood, L.J. (1987b) The phenotypic frequencies of group specific component and α-2-HS-glycoprotein in three ethnic groups. The use of these proteins as racial markers in forensic biology. Forensic Sci. Int. 35, 197-207

Whitehead, A.S., Bruns, G.A.P., Markham, A.F., Colten, H.R., & Woods, D.E. (1983) Isolation of human C-reactive protein complementary DNA and localization of the gene to chromosome 1. Science 221, 69-71

Whitham, S.E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B.J., Lyons, A., Harris, T.J.R., Reynolds, J.J., Herrlich, P., & Docherty, A.J.P. (1986) Comparison of human stromelysin and collagenase by cloning and sequence analysis. Biochem. J. 240, 913-916

Wilkinson, J.M. (1980) Troponin C from rabbit slow skeletal and cardiac muscle is the product of a single gene. Eur. J. Biochem. 103, 179-188

Wilson, J.M., Ashton, B., & Triffitt, J.T. (1977) The interaction of a component of bone organic matrix with the mineral phase. Calcif. Tiss. Res. 22, 458-460

Won, K.-A., & Baumann, H. (1990) The cytokine response element of the rat α_1-acid glycoprotein gene is a complex of several interacting regulatory sequences. Mol. Cell. Biol. 10, 3965-3978

Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V., & Ueda, H. (1987) Purification and properties of *Drosophila* heat shock activator protein. Science 238, 1247-1253

Yachnin, S. (1975) Fetuin, an inhibitor of lymphocyte transformation. The interaction of fetuin with phytomitogens and a possible role for fetuin in fetal development. J. Exp. Med. 141, 242-256

Yang, F., Brune, J.L., Baldwin, W.D., Barnett, D.R., & Bowman, B.H. (1983) Identification and characterization of human haptoglobin cDNA. Proc. Natl. Acad. Sci. USA 80, 5875-5879

Yang, F., Brune, J.L., Naylor, S.L., Cupples, R.L., Naberhaus, K.H., & Bowman, B.H. (1985a) Human group-specific component (Gc) is a member of the albumin family. Proc. Natl. Acad. Sci. USA 82, 7994-7998

Yang, F., Chen, Z.-L., Cupples, R.L., & Friedrichs, W.E. (1991) J. Biol. Chem. submitted.

Yang, F., Lum, J.B., McGill, J.R., Moore, C.M., Naylor, S.L., van Bragt P.H., Baldwin, W.D., & Bowman, B.H. (1984) Human transferrin: cDNA characterization and chromosomal localization. Proc. Natl. Acad. Sci. USA 81, 2752-2756

Yang, F., Luna, V.J., McAnelly, R.D., Naberhaus, K.H., Cupples, R.L., & Bowman, B.H. (1985b) Evolutionary and structural relationships among the group-specific component, albumin and α-fetoprotein. Nucleic Acids Res. 13, 8007-8017

Yang, F., Naylor, S.L., Lum, J.B., Cutshaw, S., McCombs, J.L., Naberhaus, K.H., McGill, J.R., Adrian, G.S., Moore, C.M., Barnett, D.R., & Bowman, B.H. (1986) Characterization, mapping, and expression of the human ceruloplasmin gene. Proc. Natl. Acad. Sci. USA 83, 3257-3261

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33, 103-119

Yet, M.-G., Chin, C.C.Q., & Wold, F. (1988) The covalent structure of individual *N*-linked glycopeptides from ovomucoid and asialofetuin. J. Biol. Chem. 263, 111-117

Yoshioka, Y., Gejyo, F., Marti, T., Rickli, E.E., Bürgi, W., Offner, G.D., Troxler, R.F., & Schmid, K. (1986) The complete amino acid sequence of the A-chain of human plasma α_2HS-glycoprotein. J. Biol. Chem. 261, 1665-1676

Yuasa, I., Taira, T., Suenaga, K., Ito, K., & Okada, K. (1985) Determination of α_2HS-glycoprotein phenotypes by isoelectric focusing and immunoblotting: polymorphic occurrence of *HSGA*5* in Okinawa. Hum. Genet. 70, 32-34

Zaitsu, H., & Serrero, G. (1990) Pedersen fetuin contains three adipogenic factors with distinct biochemical characteristics. J. Cell. Physiol. 144, 485-491

Zenke, M., Grundström, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A., & Chambon, P. (1986) Multiple sequence motifs are involved in SV40 enhancer function. EMBO J. 5, 387-397

Zervos, P.H., Morris, L.M., & Hellwig, R.J. (1988) A novel method for rapid isolation of plasmid DNA. BioTechniques 6, 238-242

****