The Complete cDNA and Amino Acid Sequence of Bovine Fetuin

ITS HOMOLOGY WITH α2HS GLYCOPROTEIN AND RELATION TO OTHER MEMBERS OF THE CYSTATIN SUPERFAMILY


From the Wessex Area Neurosciences Group and Department of Physiology and Pharmacology, University of Southampton, Bassett Crescent East, Southampton, Hants S09 3TJ, Great Britain and the Department of Biochemistry, University of Auckland, Private Bag, Auckland, New Zealand

The cDNA sequence encoding the bovine fetal protein fetuin is reported. The deduced amino acid sequence is identical with that obtained from amino acid sequencing. The protein is a single chain preceded by a signal sequence. The three N-linked glycosylation sites have been determined. The sequence of fetuin shows over 70% similarity to human α2HS glycoprotein. All of the cysteine residues are conserved in both proteins, suggesting that fetuin has the same arrangement of disulfide loops as α2HS glycoprotein and may also be a member of the cystatin family. Southern blot analysis indicates that a single gene codes for fetuin. No evidence for a separate gene for a bovine α2HS glycoprotein was obtained; thus, fetuin in cattle and α2HS glycoprotein in the human are equivalent proteins.

Fetuin from bovine serum was the first fetal protein to be described (1). It has since been documented to be present in many other species particularly those of the order Artiodactyla (e.g. sheep, pigs, cows) (2). Its function is not known. Some preparations of bovine fetuin showed strong trypsin inhibition (3), while others did not (4). It was also suggested that the antiproteolytic action of bovine fetuin may contribute to its biological effectiveness in promoting cell growth in vitro (5), as fetuin is a widely used component of tissue culture media (6). Bovine fetuin was suggested to have lymphocyte-stimulating properties (7) and lipid binding capability (8). Sheep and bovine fetuins have also been shown to have some thyroid hormone binding properties (9).

Morphological studies of developing brains demonstrated the presence of fetuin or a fetuin-like protein within neurons of the early cortical plate cells in the neocortex of such diverse animal species as sheep, cow, pig (10, 11), marsupials (12), and humans (13). In all species examined, developmentally regulated expression of fetuin-like protein followed a similar pattern, being one of the first markers of earliest cortical plate cells. Recently, a striking amino acid homology between fetuin and the human serum protein α2HS glycoprotein has been described (14). Nevertheless, several differences have been noted, particularly that antibodies to one protein do not easily cross-react with the other (13). α2HS glycoprotein is present in human plasma as two polypeptide chains (A and B) held together by a single interchain disulfide bridge (15). Both chains are encoded by a single mRNA transcript which also codes for a connecting sequence partially cleaved before the protein is released into the circulation (16, 17). In contrast, fetuin has been described in plasma as a single polypeptide chain (14, 18), and no information has been available about its possible biosynthetic precursor. The biological functions of α2HS glycoprotein and fetuin are not known. Nevertheless, structural similarities of both proteins to the cysteine proteinase inhibitors (17, 19) suggest that they may also belong to the cystatin superfamily. The aim of the present study was to investigate the structural homology between bovine fetuin and human α2HS glycoprotein and their relationship to other members of the cystatin family.

MATERIALS AND METHODS AND RESULTS

A computer-aided search of the Swisspro sequence data base revealed that the only protein homologous to fetuin was the human plasma protein α2HS glycoprotein. The deduced amino acid sequence of fetuin is compared to the published sequence of α2HS (16) in Fig. 1. Both proteins contain an 18-residue leader sequence and show a high degree of sequence identity (218 out of 341 residues in the mature proteins). This alignment is based on an insertion at position 220 and gaps of 6 and 3 residues (positions 285-286 and 323-324) in the fetuin sequence compared to α2HS. Thus, the overall sequence identity between the two proteins is 64%.

Fetuin is present in serum as a single-chain species compared to the two-chain form of α2HS. α2HS is known to be synthesized from a single precursor and processed to give a large amino-terminal A-chain linked by a disulfide bond to the 27-amino acid residue B-chain (16). This results from an initial cleavage at an Arg-Thr bond (residues 322-324). The equivalent sequence in fetuin is Pro-Ile, accounting for the existence of fetuin as a single chain species. The sequence of the α2HS A-chain previously reported was missing a 40-residue segment at the carboxyl-terminal end (20), and a secondary cleavage at a Leu-Leu bond was suggested to account for the removal of this segment, termed the connecting...
peptide (16). However, in other preparations of α2HS, the A-chain has been shown to contain the majority of this sequence (see sites of proteolysis in Fig. 1 and Ref. 17).

Bovine fetuin purified from plasma has frequently been used as a model glycoprotein for studies concerned with the determination of carbohydrate structures. The protein bears three mucin-type carbohydrate units (22, 23). By comparing the experimentally determined amino acid sequence with that determined from the cDNA, it was possible to identify positions 81, 138, and 158 as sites for N-glycosylation. Sites at 138 and 158 were also present in the α2HS A-chain, but position 81 seems to be an additional and unique glycosylation site in fetuin. These data are in agreement with a recent paper in which three N-glycosylated peptides were isolated from asialofetuin and characterized by peptide sequencing and mass spectrometry (23). No definitive data were obtained regarding the sites of O-linked carbohydrate. The peptide environment around the O-glycosylation sites of fetuin is particularly proline-rich, and the sequence Gly-Pro-(Ser-CHO)-Pro-Thr-Ala has been proposed as a site of O-glycosylation (24). A possible candidate for this site is the hexapeptide Gly-Pro-Pro-Thr-Gly-Pro-Ser found at 260–265 in the fetuin sequence. Other potential sites are at positions 253 and 323, both are conserved surrounding sequences with cystatin proteinase inhibitors (17, 19). The arrangement of the disulfide loops in α2HS is now known (17). Fetuin also contains 12 cysteine residues which are found in positions identical with those in α2HS. While the disulfide loops have not been established for fetuin, the high degree of overall sequence identity in the regions adjacent to the cysteines in α2HS indicates that the same disulfide loop structures will be present. Both α2HS and fetuin contain a protease-sensitive site. Limited proteolysis with trypsin generated a 33-kDa fragment from α2HS with the amino-terminal sequence Lys-Val-Xaa-Gln-Asp (17). In the present work, limited proteolysis of fetuin with trypsin generated a 30-kDa fragment which had an amino-terminal sequence Lys-Leu-Cys-Pro-Asp. This indicates that limited proteolysis with trypsin results in the cleavage of an Arg-Lys bond (position 195–196 of fetuin) in both α2HS and fetuin. This bond is located near the putative type C loop in fetuin (Cys-128 to Cys-131). It has been something of a puzzle that, in spite of the high degree of sequence homology between fetuin and α2HS, the cross-reactivity between specific antisera is very weak. However, antisera raised against fetuin subjected to limited proteolysis with trypsin showed reaction with both α2HS (14) and a related glycoprotein in wallaby plasma (12). It is likely that polyclonal antisera contain antibodies specific to epitopes revealed by removal of the first cystatin-like domain.

While the cystatin supergene family clearly consists of a group of proteins containing tandem-like arrays of cystatin-like domains, it appears unlikely that the function of all of these proteins is similar. For example, unlike the cystatins and kininogens, no protease inhibitory function has been proven for α2HS (17). The reactive site of cystatin-related protease inhibitors is considered to be the sequence Glu-Xaa-Val-Xaa-Gly, most commonly Gln-Val-Val-Ala-Gly which is located amino-terminal to the type A and type B disulfide loops (19, 25). The unusual but related sequence Glu-Gln-Pro-Ser-Gly is found in α2HS, while in fetuin Arg-A arg-Pro-Thr-Gly (position 53–57) is quite different, suggesting that fetuin does not function as an inhibitor of thiol protease.

It is noteworthy that the greatest variation is found in the carboxyl-terminal regions of all these proteins, and it is possible that these regions may be involved in specific (different) functions. The region of fetuin showing the greatest difference from α2HS also lies toward the carboxyl-terminal segment of fetuin, which corresponds to the carboxyl-terminal end of the A-chain of α2HS together with a part of the connecting peptide. In addition to substitutions preventing processing of fetuin to a 2-chain form (discussed above), there are a number of substitutions which increase the hydrophobicity and decrease the proline content of fetuin compared with α2HS. Only a single collagen-like sequence is found in this region compared with 3 in α2HS (20), and the sequence Pro-Pro-Gly which is a candidate site for proline hydroxylase (17) is not found in fetuin. Algorithms used to predict membrane-associated helices (26) recognize the sequence 226 to 285 found within this hydrophobic region as a potential transmembrane domain. In view of this, the role of highly charged O-linked oligosaccharide units around residue 260–265 might be to prevent aggregation and maintain solubility of the secreted protein.

The high degree of homology between fetuin and α2HS glycoprotein is intriguing, as no common function has yet been assigned to this group of proteins.
been attributed to these proteins. Fetuin in sheep and cow brains (11, 27), α2HS in human brain (10), and a related glycoprotein in tammar wallaby brain (12) are expressed in the same neuronal population at a similar developmental stage; recently, we have demonstrated the presence of mRNA for fetuin in the cells of the early cortical plate in the fetal sheep brain. At the time when fetuin-like glycoproteins are expressed in the early cortical plate, the region on either side of it is a site of intense synaptogenesis in the developing brain (28). Thus, these glycoproteins may be members of a group of recognition molecules responsible for establishing neuronal connectivity (cf. Ref. 27). Elsewhere in the fetus, α2HS has been described to be present in developing bone in concentrations much higher than in plasma (29), and its function in bone seems to be as a modulator of resorption (30). Preliminary results reported by Lee et al. (16) suggested that mRNA for α2HS glycoprotein is also present in rat embryonal cartilage.

Acknowledgements—We would like to thank Drs. Ian Clarke and Paul Lambden for invaluable advice. We are also indebted to Dr. Mark Pickett for oligonucleotide synthesis. We should also like to thank Lynn Ford and Mary Eldridge for their patient assistance in the preparation of the manuscript.

REFERENCES

Additional references are found on p. 4357.

Preparation of samples

Preparation of fetuin and α2HS was essentially as described previously (15, 16) with the exception that the material was not purified on Sepharose 4B. The samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and either silver stained or Western blotted as described previously (15, 16). The samples were run on 12% SDS-PAGE gels and stained with Coomassie brilliant blue R-250. The proteins were transferred to nitrocellulose (Hybond, Amersham) using the Transphor (Schleicher & Schuell, Dassel, Germany). The filters were probed with the monoclonal antibody 12G11, which recognizes a common epitope in the fetal liver and skeletal muscle isoforms of fetuin. The antibody was used at a concentration of 1:1000. The filter was developed using the ECL kit (Amersham).
Complete cDNA and Amino Acid Sequence of Bovine Fetuin

The majority of the fetuin coding sequence was also confirmed by amino acid sequence analysis. The sequence of a number of peptide fragments, which span the entire coding region of the cDNA sequence are presented in Table I. There is complete agreement with the cDNA sequence except for gaps corresponding to positions 81, 130, and 148. In peptides 81-130 and 148-192, insertions of one and two amino acids were observed. The insertions were also found in the cDNA at these positions, when compared to the bovine-Thy1 sequence in the corresponding regions of the cDNA sequence identifies three sites as the location of histidine carboxylate.

Southern blot analysis was performed in order to determine the number of copies of the fetuin gene in the bovine genome (Fig. 3). 5 μg of bovine genomic DNA was digested with Bgl II or EcoRI and the fragments analyzed by hybridization to a 32P DNA fragment from the 3′ end of the fetuin clone described in Fig. 2. As shown in Fig. 3, both Bgl II and EcoRI digestion gave single DNA fragments of 12 kb and 1.4 kb, respectively. This result products a single band for both restriction enzymes. The DNA fragments obtained in these experiments are consistent with the presence of a single copy of the fetuin gene in the bovine genome. The data suggest that the separate clone of the bovine fetuin was used as a probe in this experiment is very similar to the corresponding region in human aFG (12). This fetuin itself appears to be the bovine counterpart of aFG in the human.


Table I: Amino acid sequence analysis of fetuin peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>Corresponding position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetuin</td>
<td>ITLPGGTLK</td>
<td>1-20</td>
</tr>
<tr>
<td>V8</td>
<td>OAVGVRNLQ</td>
<td>21-44</td>
</tr>
<tr>
<td>13-34</td>
<td>BPGYTVPSNLCL</td>
<td>45-74</td>
</tr>
<tr>
<td>V8-13</td>
<td>TIQVPLQHCS</td>
<td>75-88</td>
</tr>
<tr>
<td>13-12</td>
<td>QTVQKLQSG</td>
<td>89-98</td>
</tr>
<tr>
<td>LT30K</td>
<td>TQTVQKLQSG</td>
<td>109-127</td>
</tr>
<tr>
<td>L310K</td>
<td>KLVKQPLQHCS</td>
<td>128-146</td>
</tr>
<tr>
<td>LT34</td>
<td>YVVSHEQVSVQ</td>
<td>147-165</td>
</tr>
<tr>
<td>LT36</td>
<td>YVVSHEQVSVQ</td>
<td>166-184</td>
</tr>
<tr>
<td>LT37</td>
<td>AQEVQPLQHCS</td>
<td>185-203</td>
</tr>
<tr>
<td>LT32</td>
<td>VQAVQPLQHCS</td>
<td>204-222</td>
</tr>
<tr>
<td>LT36</td>
<td>YVVSHEQVSVQ</td>
<td>223-241</td>
</tr>
<tr>
<td>LT37</td>
<td>AQEVQPLQHCS</td>
<td>242-260</td>
</tr>
<tr>
<td>LT38</td>
<td>VQAVQPLQHCS</td>
<td>261-279</td>
</tr>
<tr>
<td>LT39</td>
<td>YVVSHEQVSVQ</td>
<td>280-300</td>
</tr>
<tr>
<td>LT40</td>
<td>AQEVQPLQHCS</td>
<td>301-320</td>
</tr>
</tbody>
</table>

The amino acid sequences are in the one-letter code. The N-terminal sequence of fetuin and the sequences of peptide V8-13, V8-13-12, and V8-13 were taken from a previous report (13). 1, tryptic peptides; V8, Staphylococcal aureus V8 protein; LT, peptide generated by limited trypsin cleavage; LT30K, limited trypsin cleavage product; LT34, limited trypsin cleavage product. Lysozimes were identified as S-carboxymethyllysine except for LT30K fragment which was not reduced and alkylated.

Positions where N-terminal peptides could not be identified are shown by (-). Automated sequence analysis of peptide V8 V307 (2.15 μmol), LT30 (1.3 μmol) and LT34 (400 μmol) indicated that these peptides contained potential sites of N-linked carbohydrates. Peptide V8 V307, LT30K and LT-34 gave initial yields of 60%, 78.4% and 92.1% and average repetitive yields of 87.9%, 80.4% and 80.1%, respectively.
The complete cDNA and amino acid sequence of bovine fetuin. Its homology with alpha 2HS glycoprotein and relation to other members of the cystatin superfamily.

K M Dziegielewska, W M Brown, S J Casey, D L Christie, R C Foreman, R M Hill and N R Saunders


Access the most updated version of this article at http://www.jbc.org/content/265/8/4354

Alerts:

• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/8/4354.full.html#ref-list-1