THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 46, Issue of November 15, pp. 43698–43706, 2002
Printed in U.S.A.

Localization of Carbohydrate Attachment Sites and Disulfide Bridges in *Limulus* α_2-Macroglobulin

EVIDENCE FOR TWO FORMS DIFFERING PRIMARILY IN THEIR BAIT REGION SEQUENCES*

Received for publication, August 12, 2002, and in revised form, September 5, 2002
Published, JBC Papers in Press, September 5, 2002, DOI 10.1074/jbc.M208236200

Lise B. Husted‡, Esben S. Sørensen‡, Peter B. Armstrong§¶, James P. Quigley§‖, Lene Kristensen‡, and Lars Sottrup-Jensen‡**

From the ‡Department of Molecular Biology, Science Park Division, University of Aarhus, 8000 Aarhus C , Denmark, §Marine Biological Laboratory, Woods Hole, Massachusetts 02543, the ¶Department of Molecular and Cellular Biology, University of California, Davis, California 95616, and the ‖Department of Vascular Biology, The Scripps Research Institute, La Jolla, California 92037

The primary structure determination of the dimeric invertebrate α_2-macroglobulin (α_2M) from *Limulus polyphemus* has been completed by determining its sites of glycosylation and disulfide bridge pattern. Of seven potential glycosylation sites for *N*-linked glycosylation, six (Asn^{275}, Asn^{307}, Asn^{866}, Asn^{896}, Asn^{1089}, and Asn^{1145}) carry common glucosamine-based carbohydrates groups, whereas one (Asn^{80}) carries a carbohydrate chain containing both glucosamine and galactosamine. Nine disulfide bridges, which are homologues with bridges in human α_2M, have been identified (Cys^{228}–Cys^{269}, Cys^{456}–Cys^{580}, Cys^{612}–Cys^{799}, Cys^{657}–Cys^{707}, Cys^{849}–Cys^{876}, Cys^{874}–Cys^{910}, Cys^{946}–Cys^{1328}, Cys^{1104}–Cys^{1155}, and Cys^{1362}–Cys^{1475}). In addition to these bridges, *Limulus* α_2M contains three unique bridges that connect Cys^{361} and Cys^{382}, Cys^{1370} and Cys^{1374}, respectively, and Cys^{719} in one subunit with the same residue in the other subunit of the dimer. The latter bridge forms the only interchain disulfide bridge in *Limulus* α_2M. The location of this bridge within the bait region is discussed and compared with other α-macroglobulins. Several peptides identified in the course of determining the disulfide bridge pattern provided evidence for the existence of two forms of *Limulus* α_2M. The two forms have a high degree of sequence identity, but they differ extensively in large parts of their bait regions suggesting that they have different inhibitory spectra. The two forms (*Limulus* α_2M-1 and -2) are most likely present in an ~2:1 ratio in the hemolymph of each animal, and they can be partially separated on a Mono Q column at pH 7.4 by applying a shallow gradient of NaCl.

α_2-Macroglobulin (α_2M)[1] from the American horseshoe crab, *Limulus polyphemus*, is a member of the class of proteinase-binding α-macroglobulins (αMs) present in the blood of vertebrates and invertebrates (1). αMs are glycoproteins containing ~1450 residues, and they circulate as 180-kDa monomers, 360-kDa disulfide-bridged dimers, or 720-kDa tetramers noncovalently assembled from two disulfide-bridged dimers (2, 3). Proteinase binding is initiated by one or more cleavages in an ~30–60-residue stretch near the middle of the subunit (the bait region) (4, 5). This elicits a conformational change that leads to entrapment of the proteinase (5). The bound proteinase is poorly accessible to high molecular weight substrates and inhibitors, and through rapid clearance of the αM-proteinase complex the αMs play a role in controlling the level of proteolytic activity in the blood and tissues (2). In many vertebrate species, *e.g.* man, rat, mouse, and pig, two or three related αMs have been found (3). Because their sequences differ greatly in their bait regions, each αM probably controls a particular set of proteinases, although information on this is fragmentary (2, 6).

Limulus α_2M is the most extensively studied invertebrate αM. It is a 360-kDa dimer (7, 8), and its proteinase-binding characteristics (1, 9), shape (7, 8), amino acid sequence deduced from its cDNA sequence (10), and carbohydrate composition (10) have been determined. In the set of peptides generated to initiate cDNA cloning (11, 12), it was observed that residues in several positions were at variance with those predicted from the cDNA (10), suggesting that the *Limulus* α_2M used, which was purified from pooled hemolymph, was a mixture of two or more forms. Like most other known αMs, *Limulus* α_2M contains internal thiol esters (13). When activated during proteinase complex formation, the αM thiol esters rapidly react with nucleophilic groups on the attacking proteinase and other available nucleophiles (14). This process results in efficient cross-linking of the proteinase to αMs (15). However, in *Limulus* α_2M covalent proteinase binding is insignificant as the bound proteinase can be released by denaturation (7, 16). In the case of trypsin an unusual self-cross-linking reaction within the *Limulus* α_2M dimer contributes to the tight binding of trypsin (17).

Human α_2M is the only αM for which complete information on the arrangement of its intrachain and interchain disulfide bridges and positions of Asn-based carbohydrate groups is available (18, 19). Apart from the Cys residue being part of the thiol ester site, the human α_2M subunit contains 24 Cys residues of which 22 engage in 11 intrachain bridges and 2 engage in interchain disulfide bridges thereby aligning the two subunits of the α_2M dimer in an antiparallel fashion. For vertebrate αMs of known sequence most positions of disulfide bridges can readily be predicted from the data on human α_2M.

* This work was supported by grants from the Danish Science Research Council, by a basic research grant from the University of Aarhus, by the Novo Nordic Foundation (to L. S. J.), and by National Science Foundation Grant MCB 97-26771 (to P. B. A.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed: Dept. of Molecular Biology, Science Park Division, University of Aarhus, Gustav Wieds Vej 10, 8000 Århus C, Denmark. Tel.: 45-89-42-50-51; Fax: 45-89-42-50-77; E-mail: lsj@mb.au.dk.

[1] The abbreviations used are: α_2M, α_2-macroglobulin; αM, α-macroglobulin; IAA, iodoacetamide; MESA, mercaptoethanesulfonic acid; MS, mass spectrometry; RP-HPLC, reverse phase-high pressure liquid chromatography; PTH, phenylthiohydantoin.

In addition to the Cys residue engaging in thiol ester formation, the *Limulus* α_2M subunit contains 23 Cys residues of which 18 would be expected to form nine disulfide bridges equivalent with those found in human α_2M. However, for *Limulus* α_2M the pattern of disulfide bridges is ambiguous with regard to five positions (10). The Cys residues that have no counterpart in human α_2M are located at positions 360, 381, 719, 1370, and 1434. Because no free −SH groups can be detected in native *Limulus* α_2M, they are all likely to be paired, and importantly, the two subunits of the *Limulus* α_2M dimer must be connected by an uneven number of interchain bridges in contrast to the human dimer.

From the carbohydrate composition given earlier (10) *Limulus* α_2M contains both glucosamine and galactosamine (19.4 and 2.9 residues/mol subunit, respectively). As galactosamine is not present in human α_2M and probably other mammalian αMs, it was also of interest to locate the carbohydrate groups in *Limulus* α_2M. *Limulus* α_2M contains seven candidate Asn residues for attachment of glucosamine-based glycan groups (Asn^{80}, Asn^{275}, Asn^{307}, Asn^{866}, Asn^{896}, Asn^{1089}, and Asn^{1145} (10)). In contrast, carbohydrate groups containing galactosamine are frequently bound to Ser and Thr residues which, however, cannot readily be predicted.

Here we report the determination of the complete disulfide bridge pattern of *Limulus* α_2M consisting of 11 intrachain bridges and one interchain bridge. Curiously, the single interchain bridge engages a Cys residue located in the bait region. We also report the localization of six Asn residues carrying glucosamine-based carbohydrate groups, and one Asn residue carrying a carbohydrate group containing both glucosamine and galactosamine. We further provide evidence from sequencing of a number of peptides for a second *Limulus* α_2M, the sequence of which differs from that reported earlier (10) particularly in its bait region. The two forms (*Limulus* α_2M-1 and -2) are most likely present in each animal and can be partially separated by ion exchange chromatography on Mono Q at pH 7.4.

EXPERIMENTAL PROCEDURES

Proteins and Other Materials—*Limulus* α_2M was prepared as reported previously (20). Tosyl-phenylalanine chloromethyl ketone-treated trypsin, thermolysin, angiotensin II, and adrenocorticotropic hormone, fragment 18–39, were from Sigma. *Staphylococcus aureus* proteinase was from Miles. Column materials were from Amersham Biosciences, Machery-Nagel, Shandon, The Separations Group, and LC packings. ^{14}C-Labeled iodoacetamide (IAA) was from Amersham Biosciences, and standard chemicals were from Sigma, Merck, and Rathburn.

Analytical Procedures—Amino acid analysis was performed by cation exchange using established procedures (21). Automated peptide sequencing and PTH-derivative analysis were carried out as reported (22). SDS-PAGE was performed in 10–20 and 20% slab gels using the standard Tris glycine system. Mass spectra were acquired with a Bruker BIFLEX matrix-assisted laser desorption/ionization time-of-flight instrument equipped with a 1-m flight tube, a reflector, a 337-nm nitrogen laser, and a 1-GHz digitizer. Thin film matrix surfaces were prepared using the fast evaporation technique from α-cyano-4-hydroxycinnamic acid dissolved in acetone/water (99:1) to 30 μg/μl. A 0.5-μl volume of analyte (0.1–10 pmol/μl) was deposited on the matrix surface and allowed to dry onto the crystals. Spectra were obtained by averaging 20–50 single-shot spectra and calibrated externally by using the calibration constants of angiotensin II and adrenocorticotropic hormone, fragment 18–39. Theoretical protonated masses (MH^+) were calculated using the GPMAW program (Lighthouse Data, Odense, Denmark).

Main Digest—Approximately 30 mg of *Limulus* α_2M was used as starting material for the main digest. SDS-PAGE revealed that almost complete bait region cleavage had occurred during preparation and/or storage, and the material was therefore treated at pH 8.5 with 20 mM IAA to block the thiol group appearing upon thiol ester cleavage. After removal of excess reagent by gel chromatography on Sephadex G-25 in 10% formic acid, the material was freeze-dried, redissolved in 70% formic acid, and treated with 50 mg of CNBr for 20 h. After freeze-drying the degraded *Limulus* α_2M was redissolved in 300 μl of formic acid, and 10 volumes of water was added. Addition of aliquots of 5 M NaOH was used to raise the pH of the solution, and at pH 4–5 the solution turned turbid. Precipitation appeared to be complete at pH 7–8.

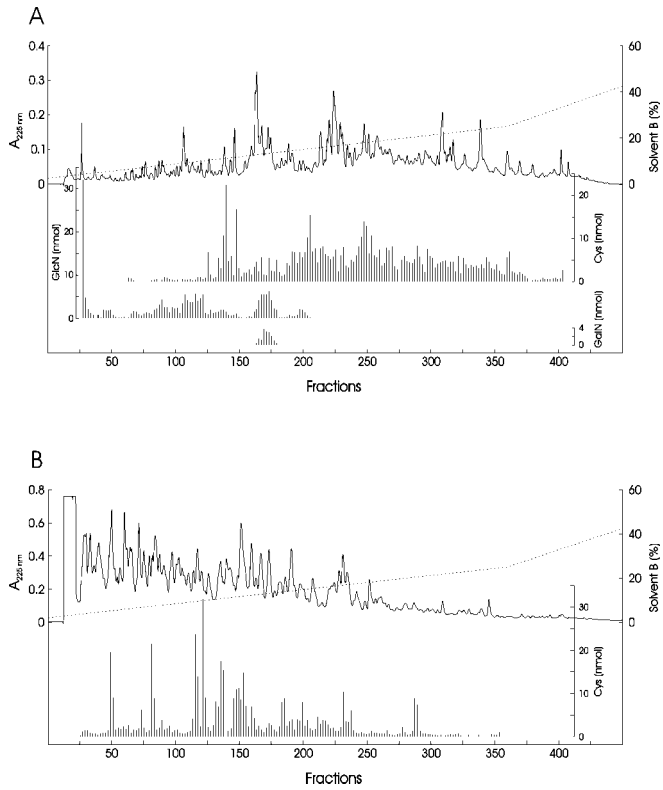

FIG. 1. **RP-HPLC separation of the two Superdex peptide pools from the main digest.** *A*, one-half of the pool containing the large disulfide-bridged peptides and all carbohydrate-containing peptides was eluted from an 8 × 250-mm Nucleosil C18 column with linear gradients formed from 0.1% trifluoroacetic acid (solvent A) and 90% acetonitrile containing 0.08% trifluoroacetic acid (solvent B) (*dashed line*). The column was operated at 50 °C at a flow rate of 2 ml/min, and 0.67-ml fractions were collected. The separation was monitored by recording the absorbance at 225 nm (*solid line*) and by determining the amount of half-cystine, glucosamine (*GlcN*), and galactosamine (*GalN*) in fractions combined in pairs (*vertical bars*). *B*, one-half of the pool containing small disulfide-bridged peptides was separated on the Nucleosil C18 column under the same conditions as in *A*. The peptides were detected at 225 nm (*solid line*), and the content of half-cystine in two successive fractions is indicated by *vertical bars*. The elution conditions and the fraction size were the same as in *A*.

In analytical experiments trypsin or thermolysin was added at pH 4, and digestion was attempted after raising the pH to 7–8. HPLC experiments showed only extensive degradation with thermolysin. Therefore, to a solution of ~18 mg of CNBr-degraded *Limulus* α_2M 1:50 (w/w) thermolysin was added at pH 4 and the pH subsequently raised to 7 by addition of Tris. After incubation for 90 min at 55 °C with stirring, the suspension had nearly cleared. To separate the larger carbohydrate-containing and disulfide peptides from the small peptides, the digest was fractionated on a Superdex peptide column using 0.1% trifluoroacetic acid, 25% acetonitrile as eluent. The column effluent was monitored by measuring the absorbance at 280 nm and by determining the amount of cysteic acid and amino sugars in each fraction after performic acid oxidation and amino acid analysis (not shown). Two pools, one containing all carbohydrate peptides and the larger disulfide peptides and one containing the small disulfide peptides, were made.

Half of each pool was separated by RP-HPLC on an 8 × 250-mm column packed with Nucleosil C18 using elution with gradients of acetonitrile in 0.1% trifluoroacetic acid (Fig. 1, *A* and *B*). Cys- and carbohydrate-containing peptides were located as above by performic acid oxidation and amino acid analysis of aliquots from each fraction combined in pairs.

Pools of interest were subjected to ion exchange chromatography on a 4 × 250-mm LC-SCX column using gradient elution with NaCl in 5

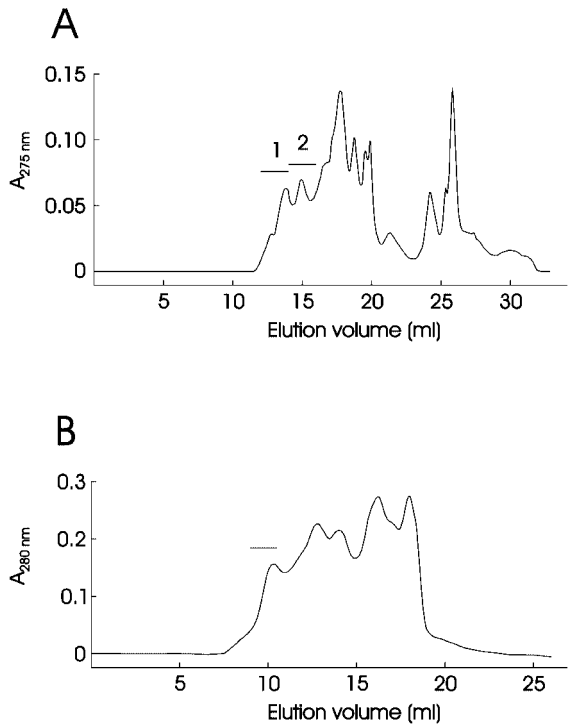

FIG. 2. **Gel chromatography of tryptic peptides from the ancillary digest on a Superdex 75 (10/30) column.** The column was equilibrated and eluted at a flow rate of 0.5 ml/min with 10 mM sodium acetate, 100 mM NaCl, pH 4.5 (*A*), and 50% formic acid (*B*). The separations were monitored by recording the absorbance at 275 (*A*) or 280 nm (*B*), and fragments of interest were identified by SDS-PAGE, amino acids, and sequence analysis. The pools indicated by *horizontal bars* in *A* and *B* were used to determine remaining glycosylation sites and disulfide bridges, respectively.

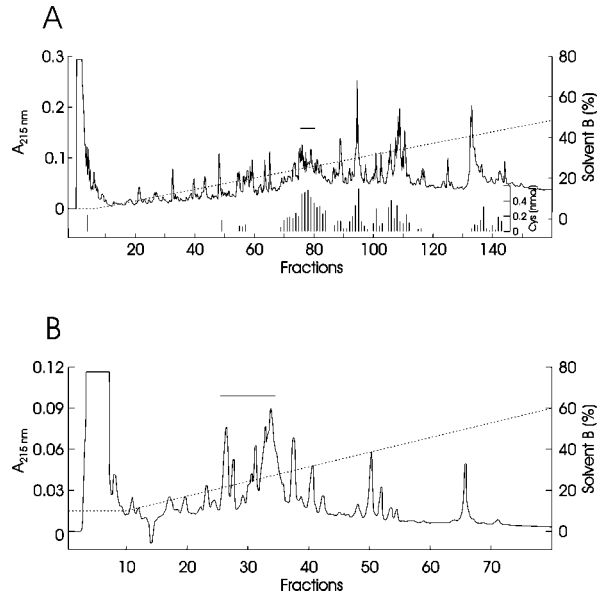

FIG. 3. **RP-HPLC separation of digests aimed at determining the remaining disulfide bridges.** *A,* an *S. aureus* proteinase digest of the pool from the Superdex 75 separation carried out in 50% formic acid (Fig. 2*B*) was separated on a 4 × 250-mm Nucleosil C18 column. The column was equilibrated with 5% solvent B (90% acetonitrile, 0.08% trifluoroacetic acid) and 95% solvent A (0.1% trifluoroacetic acid) and eluted with a gradient formed by solvent A and solvent B (*dashed line*). The separation was performed at 50 °C at a flow rate of 1 ml/min, and fractions of 0.5 ml were collected. Peptides were detected at 215 nm (*solid line*), and the amount of half-cystine in fractions having an absorbance >0.05 was determined (*vertical bars*). By MS and sequence analyses of half-cystine-containing fractions it was found that the fractions indicated by a *horizontal bar* contained a disulfide-bridged cluster involving Cys^{657}, Cys^{707}, and Cys^{719}. *B,* the fractions shown by the *horizontal bar* in *A* were digested with chymotrypsin and separated on a 2 × 250-mm Nucleosil C18 column. The column was operated at 50 °C at a flow rate of 0.2 ml/min and eluted with a gradient formed from the same solvents as in *A* (*dashed line*). The separation was monitored at 215 nm (*solid line*), and 0.1 ml fractions were collected. Fractions 26–35 (*bar*) were analyzed by MS.

mM phosphoric acid, 25% acetonitrile (23). Usually 4–10 components appeared (not shown), and Cys- and carbohydrate-containing peptides were located as above. In a number of cases ion exchange pools required a further RP-HPLC step on a 4 × 125 mm column packed with Hypersil C18 before unambiguous identification of the content could be made by amino acid and sequence analysis.

Ancillary Digest—Evidence for two disulfide bridges and the localization of carbohydrate at two potential positions were lacking from the main digest. A preparation of *Limulus* α_2M (containing intact subunits) was treated with methylamine and IAA and dialyzed into 10 mM phosphoric acid. From inspection of the sequences around the positions where peptide evidence was missing, trypsin was chosen to possibly generate a set of fairly large peptides that might relatively easily be separated from the bulk of smaller peptides. When aliquots of 1 M Tris were added to raise the pH, the solution turned turbid at pH 4–5 as seen before, followed by extensive precipitation above pH 7. However, SDS-PAGE consistently showed the generation of a set of species between ~15 and 28 kDa when 1:50 (w/w) trypsin was added at pH 4 and digestion allowed to proceed at 37 °C for 24 h at pH 6.5 (not shown). Upon electroblotting to polyvinylidene difluoride and sequence analysis, the relevant peptides were found to be present in the 15–28-kDa species of the digest.

On a preparative scale 4-mg portions of *Limulus* α_2M digested with trypsin as above were subjected to gel chromatography on Superdex 75. In one experiment the separation was done using 10 mM sodium acetate, 100 mM NaCl, pH 4.5 (Fig. 2*A*). Although a major part of the larger species of the unfractionated digest was lost by adsorption to the column matrix and could only be recovered by subsequent elution with 6 M guanidinium chloride, the soluble material upon subsequent digestion of the materials in pools 1 and 2 with *S. aureus* proteinase and RP-HPLC separation (not shown) provided peptides giving evidence for the position of the two remaining carbohydrate groups.

The material eluted with guanidinium chloride as described above was desalted and subjected to SDS-PAGE. Upon electroblotting and sequence analysis a major 26–28-kDa doublet species containing two N termini was found to cover the stretches where the missing two disulfide bridges must be located. However, in subsequent manipulations the material was lost. An additional preparative tryptic digest as above was then subjected to gel chromatography on Superdex 75 using 50% formic as solvent (Fig. 2*B*). The relevant material eluted in the void volume (*horizontal bar* in Fig. 2*B*), together with other fragments, and was digested with 1:20 (w/w) *S. aureus* proteinase at pH 7 for 2 h. RP-HPLC separation showed fair digestion (Fig. 3*A*). Upon subsequent digestion of one particular fragment set (*horizontal bar* in Fig. 3*A*) with chymotrypsin and RP-HPLC separation (Fig. 3*B*), unambiguous assignment of the two disulfide bridges present could be made from peptides present in the section of the chromatogram labeled with a *horizontal bar*.

In most cases the evidence for assigning the disulfide bridges was based on the amino acid composition of performic acid-oxidized relatively short pure peptides and sequence analysis of intact peptides. No peptides contained an internal disulfide bridge, and two sequences in near equimolar yield were seen, occasionally on a background of several minor irrelevant components. In cases where Cys_2 was released after less than 5–8 cycles of Edman degradation, bis-PTH-Cys_2 was normally seen in the RP-HPLC analysis of the PTH-derivatives eluting as a low yield peak near PTH-Tyr (24); after more than 8 cycles no signal was observed. When performed, mass spectrometry (MS) confirmed the assignment based on sequence analysis. However, in two cases MS provided the full evidence for assignment.

The evidence for locating carbohydrate groups to particular positions was based on the presence of glucosamine (and in one case also galactosamine) in hydrolysates coupled with the lack of a PTH-derivative when encountering Asn residues located in the sequence Asn-Xaa-Ser/Thr (Xaa not Pro).

Preparation and Digestion of Partially Reduced Methylamine-reacted Limulus α_2M—Prior to this experiment *Limulus* α_2M was treated with methylamine and IAA. By using 8 mM mercaptoethanesulfonic acid (MESA) at pH 8.0 for 20 min, the interchain disulfide bridge(s) in the *Limulus* α_2M dimer could be reduced to an extent of >90% as evaluated

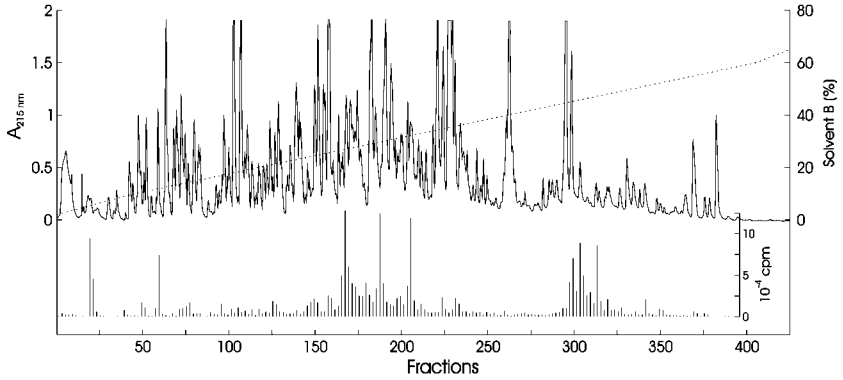

FIG. 4. **RP-HPLC elution profile of tryptic peptides generated from ^{14}C-carboxamidomethylated partially reduced methylamine-treated *Limulus* α_2M.** The peptides were separated on an 8 × 250-mm Nucleosil C18 column at 50 °C at a flow rate of 2 ml/min using gradient elution with linear gradients (*dashed line*) formed from 0.1% trifluoroacetic acid (solvent A) and 90% acetonitrile containing 0.08% trifluoroacetic acid (solvent B). The peptides were detected at 215 nm (*solid line*), and the amount of radioactivity in two successive fractions each having a size of 0.67 ml is indicated by *vertical bars*.

by non-reducing SDS-PAGE (not shown). Methylamine-reacted *Limulus* α_2M (30 mg, 4 mg/ml in 100 mM Tris-HCl, pH 8.0) reduced with MESA was freed of MESA by gel chromatography on a column of Sephadex G-25 equilibrated and eluted with degassed 20 mM sodium acetate, 100 mM NaCl, pH 4.5. The pH of the *Limulus* α_2M solution was then raised to 9.0 by addition of Tris, and 50 μCi of ^{14}C-labeled IAA was added. After reacting for 30 min the solution was made 20 mM in unlabeled IAA and allowed to react for a further 30 min. Then the excess reagents were separated from *Limulus* α_2M by a second round of gel chromatography using 20 mM Tris-HCl, 100 mM NaCl, pH 8.0, as eluent, and the protein was subsequently freeze-dried.

The labeled protein was redissolved in 2 ml of 6 M guanidinium chloride; the pH was adjusted to 9.0 by addition of Tris, and dithioerythritol was added to 10 mM to fully reduce *Limulus* α_2M. After reduction for 30 min IAA was added to 30 mM, and after 30 min of reaction the reduced and carboxamidomethylated protein was recovered by gel chromatography on Sephadex G-25 using 20 mM Tris-HCl, 100 mM NaCl, pH 8.0 as eluent. The preparation was digested with 1:50 (w/w) trypsin for 3 h at 37 °C, and after acidification with trifluoroacetic acid the digest was separated by RP-HPLC on an 8 × 250-mm column packed with Nucleosil C18 using gradient elution with acetonitrile in 0.1% trifluoroacetic acid (Fig. 4). Then the fractions containing radioactivity was located by scintillation counting, and the major peptides were further purified by RP-HPLC on a column of Hypersil C18 or by cation exchange chromatography as above (not shown).

Determination of the Amino Acid Sequence of the Bait Region of Limulus α_2M-2—One mg of *Limulus* α_2M was treated with ~2 mg of CNBr in 70% formic acid for 20 h at room temperature. After drying the degraded material was redissolved in 500 μl of 6 M guanidinium chloride, 50 mM Tris-HCl, pH 9.0, and reduced for 30 min with 10 mM dithiothreitol. The peptide solution was then acidified with trifluoroacetic acid and loaded on a 4.6 × 250-mm Vydac C4 column equilibrated with 4.5% 2-propanol in 0.1% trifluoroacetic acid. The column was eluted with a gradient of 2-propanol in 0.1% trifluoroacetic acid (Fig. 5), and the bait region peptides of ~13 and 18 kDa corresponding to each form of *Limulus* α_2M were identified by SDS-PAGE and N-terminal sequencing of samples blotted onto polyvinylidene difluoride membranes. From these experiments a 32-residue stretch containing the bait region of *Limulus* α_2M-2 was determined.

Isolation of the Two Forms of Limulus α_2M from Single Animal Hemolymph—Hemolymph (80- and 100-ml samples) was separately drawn from two animals and processed as described earlier (20). Five mg of material depleted in hemocyanin and pentraxin was subjected to gel chromatography on a Superose 6 HR 10/30 column equilibrated and eluted with 50 mM Tris-HCl, pH 7.4. The *Limulus* α_2M containing fractions near the void volume of the column were pooled and loaded on a Mono Q HR 5/5 column equilibrated with the above Tris buffer and eluted with a shallow gradient of NaCl. Two partially separated peaks appeared at [NaCl] = 190 and 220 mM, representing *Limulus* α_2M-2 and -1, respectively (Fig. 6).

RESULTS

Assignment of Disulfide Bridges and Location of Carbohydrate Groups—When determining the sequence of *Limulus* α_2M by a combination of peptide sequencing and cDNA cloning (10), we found a number of positions in which residues deviating from those determined from the cDNA sequence were present. As detailed below additional partial peptide sequence information has been obtained that does not conform with the

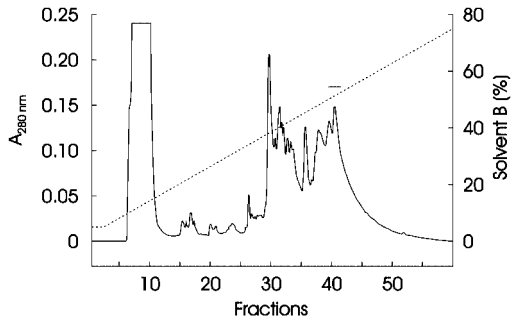

FIG. 5. **RP-HPLC separation of reduced CNBr fragments of *Limulus* α_2M on a Vydac C4 column.** The column was eluted at 50 °C with a stepwise linear gradient (*dashed line*) formed from 0.1% trifluoroacetic acid (solvent A) and 90% 2-propanol, 0.1% trifluoroacetic acid (solvent B). The flow rate was 0.75 ml/min and the fraction size was 0.5 ml. The elution was monitored by recording the absorbance at 280 nm (*solid line*), and bait region containing fragments were identified by SDS-PAGE and sequence analysis (*bar*).

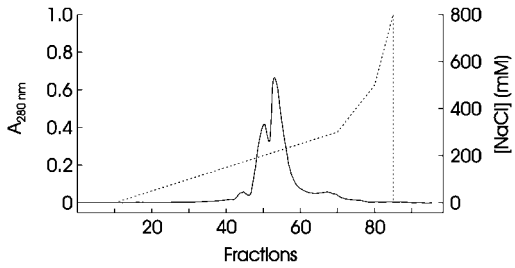

FIG. 6. **Separation of *Limulus* α_2M-1 *and Limulus* α_2M-2 from hemolymph of a single animal.** A pool obtained from gel chromatography on Superose 6 HR 10/30 was subjected to ion exchange chromatography on Mono Q HR5/5. The column was equilibrated with 50 mM Tris-HCl, pH 7.4, and eluted at a flow rate of 0.80 ml/min with a gradient of NaCl in the same buffer. Gradient breakpoints: 0 mM NaCl at 10 min, 300 mM NaCl at 67 min, 500 mM NaCl at 80 min, 700 mM NaCl at 85 min (*dashed line*). The separation was monitored by recording the absorbance at 280 nm. *Limulus* α_2M-2 eluted in fractions ~47–52 *and Limulus* α_2M-1 eluted in fractions ~50–58. The identification of the forms was made by separately degrading material from fractions 49 + 50 and 55 + 56 with CNBr followed by sequence analysis of the characteristic bait region fragments of 15 (*Limulus* α_2M-2) and 18 kDa (*Limulus* α_2M-1).

sequence deduced from the cDNA sequence. In fact, the preparation of *Limulus* α_2M used for analysis is a mixture of two closely related forms, α_2M-1 representing the published sequence (10) and α_2M-2 representing the set of deviating partial sequences.

In order to localize the glycosylation sites and disulfide bridges in *Limulus* α_2M, a digest of *Limulus* α_2M termed the main digest was made. First *Limulus* α_2M, which had been treated with IAA, was degraded with CNBr with the aim of generating a limited number of relatively large peptide clusters

TABLE I
Summary of evidence for location of carbohydrate groups to Asn residues
The pools named with prefix A originate from the separation shown in Fig. 1*A*.

Position of glycosylated Asn-residue	Peptide/peptide set	Present in pool	Variants
80	LYANGSYSSPSSNDFFFE[a]		LYANGSY . . . (A162–180)
275	VTVNS	A100	VTVNSSA . . . (A109–120)
307	NATDS	A26–28	IKMNATDS (A46–49)
866	GEKVPII . . . (833–853) bound to MQNDTSY . . . (864–880) bound to SLGQVNL . . . (889–921)[a]		
896	VN	A26–28	SLGQVNL . . . (916–946), mate in a peptide set[a]
1089	LNETEN	A126–133	
1145	YKNET	A69	GYKNET (A74–77), AGYKNET . . . (A83–88), YKNETV (A92), LEAGYKN . . . (A146), LLEAGYK . . . (A197–198)

[a] The peptides were obtained by digestion with *S. aureus* proteinase of the large fragments present in pool 1 and 2 in the separation shown in Fig. 2*A*.

which could then be further enzymatically digested. However, in contrast to human α_2M (25), most of the material appeared in one large cluster. Subsequent digestion of the material with trypsin was unsuccessful due to limited solubility at pH 7–8, but treatment with thermolysin brought most of the material into solution. Upon gel chromatography on a Superdex peptide column, the effluent was divided into two pools, one consisting of the large peptides including those containing carbohydrate (pool A) and one consisting of the small peptides (pool B) (see "Experimental Procedures"). The material in each pool was subjected to RP-HPLC (Fig. 1, *A* and *B*). In a few cases the material in the fractions obtained after RP-HPLC was of sufficient purity to allow assignment of the disulfide bridges and to locate the sites of carbohydrate attachment by compositional and sequence analysis. In general, the RP-HPLC pools made on the basis of their content of half-cystine and amino sugars were further separated by cation exchange chromatography (not shown).

Because no peptide material could be isolated from the main digest in sufficient purity to assign two particular disulfide bridges, and because evidence for localizing two carbohydrate groups was ambiguous, an ancillary digest was investigated. By using carefully controlled conditions, tryptic digestion of *Limulus* α_2M produced material which upon separation (Fig. 2, *A* and *B*) and further digestion provided the relevant peptides to complete the carbohydrate localization and bridge assignment.

The evidence for the location of carbohydrate groups to Asn residues is summarized in Table I. As seen in Fig. 1*A* glucosamine was present in many fractions in the elution profile, whereas galactosamine was only present in fractions 163–180. The major component in these fractions was a peptide with the N-terminal sequence LYANGS (77–82) apparently containing both glucosamine and galactosamine. In subjecting the material in these fractions to lectin affinity chromatography on concanavalin A and Jacalin-Sepharose to possibly separate the glucosamine- and galactosamine-containing peptides, the material was lost. However, from subdigestion of material from the ancillary digest (*pool 2* in Fig. 2*A*) the peptide LYANGSYSSPSSNDFFFE containing both amino sugars was isolated. Upon chymotryptic digestion of this peptide, it was established by amino acid analysis that the peptide LYANGSY contained both amino sugars, whereas the peptide SSPSSNDF contained none. Sequence analysis of the former peptide revealed that no PTH-Asn was present in cycle 4 and that PTH-Ser was present in normal amounts in step 6. Hence, it can be concluded that Asn^{80} carries a glycan containing both glucosamine and galactosamine.

It was further established from sequence analysis of fractions obtained from the RP-HPLC separation of pool A from the main digest (Fig. 1*A*) that Asn^{275}, Asn^{307}, Asn^{896}, Asn^{1089}, and Asn^{1145} all carry glucosamine-based carbohydrate groups. The composition of the material found in pool A26–A28 indicated that a short peptide containing Asn^{866} was present. However, no evidence from sequence analysis for carbohydrate on Asn^{866} could be obtained possibly due to cyclization of Gln^{865} during treatment with CNBr in formic acid. Upon subdigestion of the material from pool 1 in Fig. 2*A* originating from the ancillary digest, a disulfide-bridged peptide cluster composed of residues 833–853, 864–880, and 889–921 was obtained (*pool 1* in Fig. 2*A*). Sequence analysis of this peptide provided evidence for carbohydrate on Asn^{866} as well as confirmatory evidence for carbohydrate on Asn^{896}.

The sequences of peptides containing the glycosylated Asn residues shown in Table I all represent α_2M-1, and no evidence for peptides containing Asn residues in different sequences was obtained. This indicates that α_2M-2 has the same pattern of glycosylated Asn residues as α_2M-1.

In Table II the evidence for the assignment of all disulfide bridges in *Limulus* α_2M is summarized. The bridges Cys^{228}–Cys^{269}, Cys^{361}–Cys^{382}, Cys^{456}–Cys^{580}, Cys^{612}–Cys^{799}, Cys^{849}–Cys^{876}, Cys^{874}–Cys^{910}, Cys^{946}–Cys^{1328}, Cys^{1104}–Cys^{1155}, Cys^{1362}–Cys^{1475}, and Cys^{1370}–Cys^{1434} were all identified from the main digest. Most of the peptides containing these disulfide bridges were recovered as several cleavage variants, some of which are listed in Table II. With the exception of the bridges considered below, the sequences of the peptide mates originated from α_2M-1. In the case of the bridges Cys^{1104}–Cys^{1155} and Cys^{1362}–Cys^{1475}, the mates containing Cys^{1155} and Cys^{1475} did not originate from α_2M-1, but rather from α_2M-2, and the stretches containing Cys^{849} and Cys^{1434} were recovered as variants originating from both α_2M-1 and α_2M-2.

Despite extensive search no peptides containing Cys^{657}, Cys^{707}, and Cys^{719} could be identified from the main digest. This could be due to precipitation of this material during the preparation of the digest. An ~28-kDa tryptic fragment, which contains the above-mentioned Cys residues, was generated in the ancillary digest. Upon electroblotting and sequencing two N termini, ^{656}YXEDYK and ^{664}QTEGEHEG were determined in equimolar yield. The fragment must contain ~250 residues, and hence Cys^{657} in the peptide YCEDYK must be disulfide-bound to Cys^{707} or Cys^{719} present in a large fragment containing the bait region. Furthermore, Cys^{707} or Cys^{719} must engage in bridge formation with the corresponding residue in another large fragment, *i.e.* the ~28-kDa fragment set is a heterotetramer.

TABLE II

Summary of evidence for assignment of disulfide bridges in Limulus α_2M

The pools indicated with prefix A and B are from the separations shown in Fig. 1*A* and Fig. 1*B*, respectively. Pools named with the prefix Ct are from the separation shown in Fig. 3*B*. For masses below 1200-Da monoisotopic masses were calculated and above 1200-Da average masses are shown. The underlined amino acid residues deviate from the amino acid sequence deduced from cDNA cloning.

Bridge position	Peptide set	Present in pool	MH^+ found	MH^+ calc.	Variants
228–269	ICAQ/IDGC	B131–138			KIC/IDGC (B149–152) QICAQ/LIDGC (A203–206)
361–382	LCR/IRACKE	B183–186			LCRF/IRACKE (A and B287–290)
456–580	IECGKP/VKKC	B121–124	1120.7	1120.6	ITEEIEC(GKP)/VKKC (B159–162)
612–799	ICG/IC	B113–118			
657–707	TRPCKPSGF/YCEDYK	Ct28	1811.31	1811.0	
719–719	EDGGRPCPQFDE/EDGGRPCPQFDE	Ct31	2697.91	2697.8	
849–876	LSDCLP/VCGGKSDT	A251–254	1410.7	1411.6	LSECLP/VCGGKSDTTRWM (A263–270)
874–910	TSC/ICGNQD	B73–76	956.7	956.3	TSC/ICGNQDY (B145–148) TSC/ICGNQDYST (A146–150, B149–150)
946–1328	FACPKDQNGK/ATGSGCG	A139–144, B135–138	1658.1	1657.8	FACPKDQNGK/ATGSGCGL (A219–222)
1104–1155	SNGC/VRC	B49–52	754.4	754.3	
1362–1475	LYRDCNN/IDENCEKLPP	A239–246	2053.3	2053.3	LYRDCNN (with unidentified C-terminal extensions)/ IDENCEKLPP (A255–258, A267–270, A299–302, A307–310)
1370–1434	IATC/LC	B153–156	639.4	639.3	IATC/LTDQKQC (A139–144)

The fragment set was subdigested with *S. aureus* proteinase after partial separation from several other fragments on a Superdex 75 column that was equilibrated and eluted in 50% formic acid (Fig. 2*B*). The *S. aureus* proteinase digest was fractionated by RP-HPLC (Fig. 3*A*), and a smaller fragment was obtained that contained the three Cys residues. This fragment was further digested with chymotrypsin, and the digest was again fractionated by RP-HPLC (Fig. 3*B*). By MS analysis of fractions from this RP-HPLC separation, it was demonstrated that Cys^{657} is connected to Cys^{707} (expected peptide set TRPCKPSGF bound to YCEDYK (α_2M-1), Table II) and that Cys^{719} is engaged in forming an interchain bridge within the *Limulus* α_2M dimer. In this case no MS data were obtained from the peptide expected from α_2M-1 (dimer of ^{713}EDGGRPCPQYDVAF), but rather from α_2M-2 as the dimeric peptide set ^{713}EDGGRPCPQFDE was identified (Table II). That sequence was known from peptides isolated from partial reduction experiments described below.

Partial Reduction of Methylamine-treated Limulus α_2M—Due to the particular location of the 24 half-cystine residues in human α_2M, whose pairing to 12 bridges was relatively easily established (18), the identification of two of these as being engaged in interchain bridges was made by partial reduction using the highly solvated MESA and radiolabeling (19). As judged from SDS-PAGE (not shown), the subunits of methylamine-reacted *Limulus* α_2M were likewise essentially fully separated using 8 mM MESA for 30 min at room temperature. Following alkylation with radiolabeled IAA and removal of excess reagents, the preparation was fully reduced and alkylated with unlabeled IAA. The material was then digested with trypsin and separated by RP-HPLC (Fig. 4). The fractions containing the major part of the radioactivity were further purified, and two sets of peptides were characterized from these fractions as summarized in Table III.

Two major radiolabeled peptides present in pools 166–168 and 187–188 (Fig. 4), respectively, originated from unexpected cleavage by trypsin at Arg-Pro (705–706). Upon sequence analysis they were found to contain the same N-terminal sequence PCKPSGFEDGGRPCPQ. The radiolabel was solely found in position Cys^{719}, hence demonstrating that the interchain disulfide bridge containing this Cys residue is indeed solvent-exposed. However, downstream of Cys^{719} the sequences diverged, with that of the peptide from pool 166–168 being identical to the sequence expected from the bait region of α_2M-1 and that of pool 187–188 being different (α_2M-2, Table III). This shows that a fairly long stretch of the bait regions of α_2M-1 and α_2M-2 differs markedly in sequence.

The cleavage of partially reduced *Limulus* α_2M by trypsin was incomplete as evidenced by the distribution of a major part of the label in several late eluting fractions. The peptide in pool 204–205, originating from *Limulus* α_2M-2, apparently was that of pool 187–188 having a long C-terminal extension which, however, could not be identified due to poor sequencing yields. Two major labeled peptides present in pools 303–305 and 312–315, respectively, had N termini reflecting cleavage at Arg-Gln (663–664). Upon subdigestion with chymotrypsin several peptides, which could not be adequately purified, indicated that they each contained the bait region and hence represented N-terminally extended versions of the bait region peptides. The specific activity of all peptides containing the bait region was ~40,000 cpm/nmol.

From two other pools peptides having an ~10-fold lower specific activity than the bait region peptides described above were isolated. These peptides were mates of the C-terminal disulfide bridge Cys^{1362}–Cys^{1475} pointing to solvent exposure of that bridge located in the part of *Limulus* α_2M presumably being equivalent with the receptor-binding domain of human α_2M (26, 27). In addition, two versions of the mate containing Cys^{1362} originating from both α_2M-1 and α_2M-2 were found (Table III).

Determination of the Complete Bait Region Sequence of Limulus α_2M-2—In order to determine the entire bait region sequence of *Limulus* α_2M-2 and to be able to distinguish between the two forms of *Limulus* α_2M, we took advantage of the fact that the partial bait region sequence of α_2M-2 contains a Met residue at position 728 not present in α_2M-1 (Table III). By assuming that α_2M-2 like α_2M-1 has a Met residue at position 864, CNBr degradation would generate ~15- and 18-kDa fragments, respectively. After incubation with CNBr the fragments generated were reduced and separated on a Vydac C4 column (Fig. 5). Both the 18-kDa fragment containing the entire bait region of α_2M-1 and the 15-kDa fragment containing the C-terminal part of the bait region of α_2M-2 were found to elute in fractions 40 and 41. By Edman degradation of samples

TABLE III
Major radiolabeled tryptic peptides isolated after partial reduction of methylamine-reacted Limulus α_2M with MESA
The underlined residues differ from the sequence deduced from cDNA cloning. The Cys residue shown in boldface was radiolabeled.

Peptide	LAM residues	Pool (Fig. 4)
PCKPSGFEDGGRP**C**PQYDVAFAAPQAANR	706–734	166–168
PCKPSGFEDGGRP**C**PQFDEVGIPMAYGASK	706–735	187–188
PCKPSGFEDGGRP**C**PQFDEVGIPMAYGASK . . .	706– . . .	204–205
QSEGEHEGSF . . .	664– . . .	303–305, 312–315
D**C**NNAR	1362–1367	20–21
D**C**INAH	1362–1367	59–60
SIDEN**C**EK	1470–1478	59–60

electroblotted onto polyvinylidene difluoride, the remaining part of the sequence of the bait region of α_2M-2 was obtained from the 15-kDa fragment. The complete sequence of the bait region of α_2M-2 is aligned with the bait region sequences of *Limulus* α_2M-1 (10) and tick αM (28) in Fig. 7.

Isolation and Characterization of the Two Forms of Limulus α_2M from Single Animal Hemolymph—By having established that *Limulus* α_2M prepared from pooled hemolymph contains two related forms, we investigated whether each form was present in individual animals. α_2M was prepared from two animals using the procedure of Ref. 20. In the final ion exchange step of each experiment (Fig. 6), two partially separated peaks appeared. Each peak represented pure and active *Limulus* α_2M as judged from the presence of the 180-kDa subunit and the heat fragments of ~55 and 125 kDa upon reducing SDS-PAGE. Samples were removed from the flanking parts of the partially separated proteins, treated with CNBr, and separated by SDS-PAGE. Electroblotting and sequence analysis as above identified the material eluting at [NaCl] = 190 mM as α_2M-2 (containing the 15-kDa bait region fragment), and the material eluting at [NaCl] = 220 mM as α_2M-1 (containing the 18-kDa bait region fragment). α_2M-2 had the same N-terminal sequence as α_2M-1 (15 residues determined). Based on the heights of the two peaks in Fig. 6, α_2M-1 and α_2M-2 were present in an approximate 2:1 molar ratio in both animals investigated.

DISCUSSION

In this work we have completed the primary structure determination of *Limulus* α_2M by determining its disulfide bridge pattern and localizing its sites of carbohydrate attachment. Furthermore, evidence for the occurrence of two forms of *Limulus* α_2M (α_2M-1 and α_2M-2) has been obtained.

It was found that all seven potential *N*-linked glycosylation sites are occupied. As shown in the schematic comparison of the localization of disulfide bridges and glycosylation sites in *Limulus* and human α_2M in Fig. 8, the only *N*-linked glycosylation site that is conserved in human α_2M is Asn^{896}. Even though *Limulus* α_2M contains galactosamine (10), we could not identify any Ser or Thr residues containing carbohydrate. Instead we found that Asn^{80} carries a glycan containing both glucosamine and galactosamine. The presence of galactosamine-containing *N*-linked carbohydrate chains has been reported previously in several other proteins including pituitary glycoprotein hormones (29), human urinary kallidinogenase (30), human tissue factor pathway inhibitor (31), bovine component PP3 (32), snake venom batroxobin (33), and hemocyanin from the freshwater snail *Lymnaea stagnalis* (34) but not in other αMs.

Regarding the disulfide bridge pattern of *Limulus* α_2M, we have confirmed the existence of nine disulfide bridges that are located in a similar position in human α_2M as shown in Fig. 8. The remaining five Cys residues in *Limulus* α_2M, which are not part of the thiol ester site, were shown to be engaged in three disulfide bridges. One of these bridges (Cys^{360}–Cys^{381}) is located in the N-terminal part of the protein. Another bridge connecting Cys^{1370} and Cys^{1434} is located in the region corresponding to the receptor-binding domain of the mammalian αMs. The location of a bridge at this position is compatible with the three-dimensional structure of the receptor-binding domain of bovine α_2M, because the side chains of the equivalent residues in bovine α_2M are located in close proximity to each other on β-strands 2 and 7, respectively (35). Furthermore, a disulfide bridge is located in a similar position in human C3 (36). The third bridge, which is unique to *Limulus* α_2M, is an interchain disulfide bridge that connects Cys^{719} in one subunit with the same residue in the other subunit of the dimer. In a previous alignment of the bait region of *Limulus* α_2M with the bait regions of other αMs, this Cys residue was placed at the N-terminal border of the bait region (10). However, this location is not in accordance with our determination of the disulfide bridge pattern, because we found that the bridge, which defines the N-terminal boundary of the bait region, is Cys^{657}–Cys^{707} (Fig. 7). This implies that the bait region is 12 residues longer than previously thought (10) and that the above-mentioned interchain disulfide bridge is located in the bait region.

To investigate whether *Limulus* α_2M contains other interchain disulfide bridges than the above-mentioned, *Limulus* α_2M was partially reduced with MESA, and the freed thiol groups were radiolabeled with IAA. Because MESA is highly solvated it preferably reduces solvent-exposed disulfide bridges. As expected, it was found that the interchain bridge linking the two bait regions was easily reduced. However, the bridge Cys^{1362}–Cys^{1475} was also reduced by MESA. Because this bridge is conserved in human α_2M and is located in the region corresponding to the receptor-binding domain of the mammalian αMs and, furthermore, is reduced only to an extent of ~10% compared with the bait region interchain bridge, we conclude that the bridge linking the bait regions is the only interchain disulfide bridge in *Limulus* α_2M.

The location of an interchain disulfide bridge within the bait region of *Limulus* α_2M shows that parts of the two bait regions of the *Limulus* α_2M dimer are located in close proximity to each other. This observation is in line with studies on recombinant bait region variants of human α_2M. By deleting parts of the C-terminal end of the bait region of human α_2M, it was shown that the bait regions are involved in forming the interface between its non-covalently associated dimers (37). This was further examined by mutating single residues in the bait region to Cys residues. These cysteine-containing variants formed disulfide-linked tetramers demonstrating that at least two bait regions are located close to each other at the interface between the non-covalently associated dimers of human α_2M (38).

The two interchain bridges in human α_2M are located in a completely different part of the primary structure than the single interchain bridge of *Limulus* α_2M, namely in the N-terminal part. This difference in location of the interchain bridges in tetrameric human and dimeric *Limulus* α_2M seems to be reflected in the functionality of the disulfide-linked dimers. Evidence suggests that the disulfide-linked dimers of

```
                  —707                                                                      758—
Limulus α2M-1     TRPC|KPSGFEDGGRPCPQYDVAFAAPQAANRIGGGGEAGGFG---GGIRKKTNKPVV---|EIR

Limulus α2M-2     TRPC|KPSGFEDGGRPCPQFDEVGIPMAYGASKVKAGASGYKSGSVGGIRKKTXKPXV---|E

Tick αM           TRPC|REAIYDRPPSAFAAPASFHTKIQPVALAQTGSVAESFDAALPGAPGPAGVPAKSAV|EVR
```

FIG. 7. **Alignment of bait region sequences from invertebrate αMs.** The bait region is defined as the peptide stretch located between residues 666–706 in human α_2M (6). In *Limulus* α_2M-1 these residues correspond to residues 707–757. The *underlined* peptide stretches differ between *Limulus* α_2M-1 and -2.

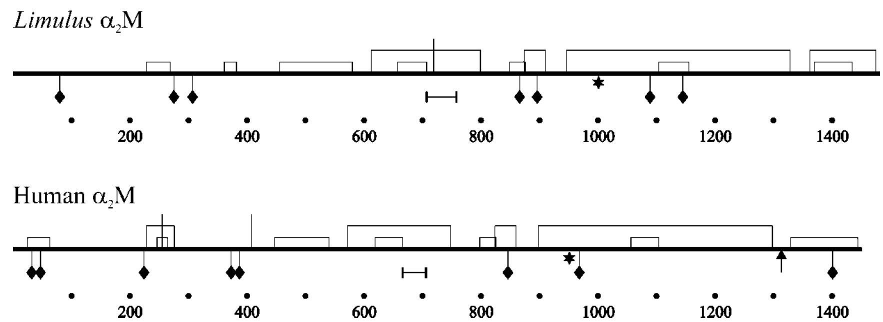

FIG. 8. **Schematic comparison of the disulfide bridge pattern and sites of glycosylation between *Limulus* and human α_2M.** The location of the 11 intrachain bridges and the single interchain bridge in *Limulus* α_2M is shown as well as the location of the 11 intrachain and 2 interchain bridges in human α_2M. The positions of carbohydrate groups are shown by *filled diamonds.* The bait regions are indicated by *horizontal bars* and the thiol esters with an *asterisk.* The C-terminal receptor-binding domain of human α_2M starts at the site shown with an *arrow.*

human α_2M are non-functional by consisting of two halves of different traps (37, 39, 40). In contrast, the *Limulus* α_2M dimer is known to be an inhibitory unit. Because the interchain bridges are located in different areas of the primary structure in *Limulus* and human α_2M, it is likely that the *Limulus* dimer resembles the non-covalently linked dimers of human α_2M, supporting the view that these dimers are the functional dimers.

By using PCR primers based on the peptide sequences ^{905}PNEAICG and ^{1210}YYWGNSI present in two of a set of peptides isolated from *Limulus* α_2M (11, 12), the cDNA and deduced protein sequence of *Limulus* α_2M was determined (10). However, upon aligning the sequences of 24 isolated peptides covering 499 residues with the deduced sequence of *Limulus* α_2M, it was observed that 23 positions differed as shown in the left part of Table IV. This suggests that the preparations from pooled hemolymph used for peptide isolation consisted of at least two related proteins. Additional evidence supporting this was provided by several peptides isolated during the determination of the disulfide bridge pattern in this study (Table IV, right part). In six cases we have isolated Cys-containing peptides that covered the same peptide stretch but differed at one or more positions within this stretch (Tables II and III). This demonstrates that two forms of *Limulus* α_2M actually exist.

As described under "Results," both forms probably have the same glycosylation and disulfide bridge pattern, but their bait region sequences differ greatly in an ~25-residue stretch (Fig. 7). A comparison of the bait region sequences of α_2M-1 and α_2M-2 with that of tick αM (28) shows that apart from a few scattered short motifs there is no extensive sequence similarity between either bait region of the two forms of *Limulus* α_2M and tick αM (Fig. 7). This is in line with data on the bait regions of vertebrate αMs that differ greatly both in length and sequence (6). The lack of a Cys residue in the bait region of tick αM is in agreement with tick αM most likely being a dimer of non-covalently associated processed subunits (28).

In contrast to the dissimilarity between the bait regions of tick αM and the two forms of *Limulus* α_2M, there is a high degree of sequence identity between the N-terminal and C-terminal parts of the bait regions of the two forms of *Limulus* α_2M. Despite this sequence identity, it is likely that the two forms are able to inhibit different proteinases because there is no similarity in the central part of the bait region. Previously, a significant degree of sequence identity within the bait region has only been observed among closely related αMs such as carp α_2Ms (41) and the two isoforms of rat α_1-inhibitor 3 (42, 43), suggesting that the two forms of *Limulus* α_2M have a high degree of sequence identity too. This is further supported by the fact that even though peptide stretches covering almost 900 amino acid residues have been sequenced only 52 differences between the two forms have been observed (Table IV). In line

TABLE IV
Identified positions of polymorphism in Limulus α_2M

Iwaki *et al.*, position		Present work	
	Residue[a]	Position	Residue[a]
235	Gln/Thr	590	Gly/Gln
292	Asn/Lys	722	Tyr/Phe
299	Glu/Ala	724	Val/Glu
609	Tyr/His	725	Ala/Val
648	Lys/Arg	726	Phe/Gly
665	Thr/Ser	727	Ala/Ile
670	Glu/Gly	728	Ala/Pro
687	Thr/Ala	729	Pro/Met
753	Asn/Glu	730	Gln/Ala
902	Ala/Thr	731	Ala/Tyr
1005	Val/Ile	732	Ala/Gly
1150	Asp/Asn	733	Asn/Ala
1153	Ile/Val	734	Arg/Ser
1154	Ser/Arg	735	Ile/Lys
1208	Gly/Ser	736	Gly/Val
1209	Arg/Pro	737	Gly/Lys
1237	Leu/Met	738	Gly/Ala
1346	Lys/Glu	740	Glu/Ala
1351	Glu/Asp	741	Ala/Ser
1352	Val/Ile	743	Gly/Tyr
1354	Val/Ala	744	Phe/Lys
1432	Asn/Lys	745A	Gly/Ser
1433	Leu/Gln	745B	–/Gly
		745C	–/Ser
		745D	–/Val
		848	Glu/Asp
		861	Lys/Thr
		1363	Ile/Asn
		1366	His/Arg

[a] The first residue corresponds to the residue obtained from cDNA cloning and the second residue to that found by sequence analysis of an isolated peptide.

with this, in the separation systems devised so far, *e.g.* in Ref. 20, the two forms were found to co-purify.

Hypothetically the two forms might be encoded by a single gene pair or by two gene pairs. To settle this question we prepared *Limulus* α_2M from the hemolymph of two animals. We obtained in both cases a partial separation of the two forms in ion exchange experiments, strongly suggesting that they are expressed by two gene pairs, similar to many other species.

REFERENCES

1. Quigley, J. P., and Armstrong, P. B. (1985) *J. Biol. Chem.* **260,** 12715–12719
2. Sottrup-Jensen, L. (1989) *J. Biol. Chem.* **264,** 11539–11542
3. Sottrup-Jensen, L. (1987) in *The Plasma Proteins* (Putnam, F. W., ed) 2nd Ed., Vol. 5, pp. 191–291, Academic Press, Orlando, FL
4. Harpel, P. C. (1973) *J. Exp. Med.* **138,** 508–521
5. Barrett, A. J., and Starkey, P. M. (1973) *Biochem. J.* **133,** 709–724
6. Sottrup-Jensen, L., Sand, O., Kristensen, L., and Fey, G. H. (1989) *J. Biol. Chem.* **264,** 15781–15789
7. Enghild, J. J., Thøgersen, I. B., Salvesen, G., Fey, G. H., Figler, N. L., Gonias, S. L., and Pizzo, S. V. (1990) *Biochemistry* **29,** 10070–10080
8. Armstrong, P. B., Mangel, W. F., Wall, J. S., Hainfield, J. F., van Holde, K. E., Ikai, A., and Quigley, J. P. (1991) *J. Biol. Chem.* **266,** 2526–2530
9. Quigley, J. P., and Armstrong, P. B. (1983) *J. Biol. Chem.* **258,** 7903–7906
10. Iwaki, D., Kawabata, S.-I., Miura, Y., Kato, A., Armstrong, P. B., Quigley, J. P., Nielsen, K. L., Dolmer, K., Sottrup-Jensen, L., and Iwanaga, S. (1996) *Eur. J. Biochem.* **242,** 822–831
11. Sottrup-Jensen, L., Borth, W., Hall, M., Quigley, J. P., and Armstrong, P. B. (1990) *Comp. Biochem. Physiol. B* **96,** 621–625
12. Armstrong, P. B., Sottrup-Jensen, L., Ikai, A., Srimal, S., and Quigley, J. P. (1994) *Ann. N. Y. Acad. Sci.* **737,** 188–201
13. Armstrong, P. B., and Quigley, J. P. (1987) *Biochem. J.* **248,** 703–707
14. Sottrup-Jensen, L., Hansen, H. F., Pedersen, H. S., and Kristensen, L. (1990) *J. Biol. Chem.* **265,** 17727–17737
15. Sottrup-Jensen, L. (1994) *Ann. N. Y. Acad. Sci.* **737,** 172–187
16. Quigley, J. P., Ikai, A., Arakawa, H., Osada, T., and Armstrong, P. B. (1991) *J. Biol. Chem.* **266,** 19426–19431
17. Dolmer, K., Husted, L. B., Armstrong, P. B., and Sottrup-Jensen, L. (1996) *FEBS Lett.* **393,** 37–40
18. Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Wierzbicki, D. M., Jones, C. M., Lønblad, P. B., Magnusson, S., and Petersen, T. E. (1984) *J. Biol. Chem.* **259,** 8318–8327
19. Jensen, P. E. H., and Sottrup-Jensen, L. (1986) *J. Biol. Chem.* **261,** 15863–15869
20. Swarnakar, S., Asokan, R., Armstrong, P. B., and Quigley, J. P. (1996) *Biochem. J.* **347,** 679–685
21. Sottrup-Jensen, L. (1993) *Biochem. Mol. Biol. Int.* **30,** 789–794
22. Sottrup-Jensen, L. (1995) *Anal. Biochem.* **225,** 187–188
23. Crimmins, D. L., Thoma, R. S., McCourt, D. W., and Schwartz, B. D. (1989) *Anal. Biochem.* **176,** 255–260
24. Bendixen, E., Halkier, T., Magnusson, S., Sottrup-Jensen, L., and Kristensen, T. (1992) *Biochemistry* **31,** 3611–3617
25. Sottrup-Jensen, L., Stepanik, T. M., Jones, C. M., Lønblad, P. B., Kristensen, T., and Wierzbicki, D. M. (1984) *J. Biol. Chem.* **259,** 8293–8303
26. Van Leuven, F., Marynen, P., Sottrup-Jensen, L., Cassiman, J.-J., and Van Den Berghe, H. (1986) *J. Biol. Chem.* **261,** 11369–11373
27. Sottrup-Jensen, L., Gliemann, J., and Van Leuven, F. (1986) *FEBS Lett.* **205,** 20–24
28. Kopácek, P., Weise, C., Saravanan, T., Vitová, K., and Grubhoffer, L. (2000) *Eur. J. Biochem.* **267,** 465–475
29. Manzella, S. M., Hooper, L. V., and Baenziger, J. U. (1996) *J. Biol. Chem.* **271,** 12117–12120
30. Tomiya, N., Awaya, J., Kurono, M., Hanzawa, H., Shimada, I., Arata, Y., Yoshida, T., and Takahashi, N. (1993) *J. Biol. Chem.* **268,** 113–126
31. Smith, P. L., Skelton, T. P., Fiete, D., Dharmesh, S. M., Beranek, M. C., MacPhail, L., Broze, G. J., Jr., and Baenziger, J. U. (1992) *J. Biol. Chem.* **267,** 19140–19146
32. Sørensen, E. S., and Petersen, T. E. (1993) *J. Dairy Res.* **60,** 535–542
33. Tanaka, N., Nakada, H., Itoh, N., Mizuno, Y., Takanishi, M., Kawasaki, T., Tate, S.-I., Inagaki, F., and Yamashina, I. (1992) *J. Biochem.* (*Tokyo*) **112,** 68–74
34. Van Kuik, J. A., Sijbesma, R. P., Kamerling, J. P., Vligenthart, J. F. G., and Wood, E. J. (1987) *Eur. J. Biochem.* **169,** 399–411
35. Jenner, L., Husted, L., Thirup, S., Sottrup-Jensen, L., and Nyborg, J. (1998) *Structure* **6,** 595–604
36. Dolmer, K., and Sottrup-Jensen, L. (1993) *FEBS Lett.* **315,** 85–90
37. Gettins, P. G. W., Hahn, K.-H., and Crews C. C. (1995) *J. Biol. Chem.* **270,** 14160–14167
38. Bowen, M. E., and Gettins, P. G. W. (1998) *J. Biol. Chem.* **273,** 1825–1831
39. Kolodziej, S. J., Schroeter, J. P., Strickland, D. K., and Stoops, J. K. (1996) *J. Struct. Biol.* **116,** 366–376
40. Qazi, U., Gettins, P. G. W., and Stoops. J. K. (1998) *J. Biol. Chem.* **273,** 8987–8993
41. Mutsuro, J., Nakao, M., Fujiki, K., and Yano, T. (2000) *Immunogenetics* **51,** 847–855
42. Braciak, T. A., Northemann, W., Hudson, G. O., Shiels, B. R., Gehring, M. R., and Fey, G. H. (1988) *J. Biol. Chem.* **263,** 3999–4012
43. Aiello, L. P., Shia, M. A., Robinson, G. S., Pilch, P. F., and Farmer, S. R. (1988) *J. Biol. Chem.* **263,** 4013–4022

Localization of Carbohydrate Attachment Sites and Disulfide Bridges in *Limulus* α_2 -Macroglobulin: EVIDENCE FOR TWO FORMS DIFFERING PRIMARILY IN THEIR BAIT REGION SEQUENCES

Lise B. Husted, Esben S. Sørensen, Peter B. Armstrong, James P. Quigley, Lene Kristensen and Lars Sottrup-Jensen

J. Biol. Chem. 2002, 277:43698-43706.
doi: 10.1074/jbc.M208236200 originally published online September 5, 2002
