(Gronenborn et al., 1991), some differences in both hydrophobic packing and amide proton exchange characteristics are apparent. To further understand the differences in thermal stability and IgG binding between B1 and B2, a more detailed structure of the B2-domain using distance geometry methods is required. This work is currently in progress.

REFERENCES

Alexander, P., Fahnestock, S., Lee, T., Orban, J., & Bryan, P. (1992) *Biochemistry* (preceding paper in this issue).

Bax, A., & Drobny, G. P. (1985) *J. Magn. Reson. 61*, 306-320.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93-106.

Deisenhofer, J. (1981) *Biochemistry 20*, 2361-2370.

Drobny, G. P., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. E. (1979) *Symp. Faraday Soc. 23*, 49-55.

Eich, G. Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3732-3733.

Fahnestock, S. R., & Alexander, P. (1990) in *Bacterial Immunoglobulin Binding Proteins* (Boyle M. P. D., Ed.) Vol. 2, pp 417-424, Academic Press, San Diego.

Fahnestock, S. R., Alexander, P., Nagle, J., & Filpula, D. (1986) *J. Bacteriol. 167*, 870-880.

Fahnestock, S. R., Alexander, P., Filpula, D., & Nagle, J. (1990) in *Bacterial Immunoglobin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 1, pp 133-148, Academic Press, San Diego.

Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., & Clore, G. M. (1991) *Science, 253*, 657-661.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546-4553.

Kumar, A., Ernst, R. R., & Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1-6.

Lian, L.-Y., Yang, J. C., Derrick, J. P., Sutcliffe, M. J., Roberts, G. C. K., Murphy, J. P., Goward, C. R., & Atkinson, T. (1991) *Biochemistry 30*, 5335-5340.

Macura, S., Huang, Y., Sater, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259-281.

Otting, G., Widmer, H., Wagner, G., & Wuthrich, K. (1985) *J. Magn. Reson. 66*, 187-193.

Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.

Torigoe, H., Shimada, I., Saito, A., Sato, M., & Arata, Y. (1990) *Biochemistry 29*, 8787-8793.

Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Complete Primary Structure of Bovine β_2-Glycoprotein I: Localization of the Disulfide Bridges[†,‡]

Emøke Bendixen, Torben Halkier,* Staffan Magnusson,[§] Lars Sottrup-Jensen, and Torsten Kristensen

Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 Aarhus C, Denmark

Received September 25, 1991; Revised Manuscript Received January 24, 1992

ABSTRACT: The complete primary structure of bovine β_2-glycoprotein I was determined by a combination of cDNA and peptide sequencing. Bovine β_2-glycoprotein I was purified from citrated plasma, and by sequencing selected peptides, the complete disulfide bridge patterns of the 11 disulfide bridges were established as well as the positions of the five asparagine-linked carbohydrate groups. Bovine β_2-glycoprotein I comprises five mutually homologous domains or Short Consensus Repeats, each containing two disulfide bridges, except for the fifth most C-terminal domain which diverges from the Short Consensus Repeat consensus by containing an additional disulfide bridge. In the four N-terminal domains, the first and third and the second and fourth half-cystines are disulfide-linked, while in the fifth domain the first and fourth, the second and fifth, and the third and sixth half-cystines are disulfide-linked.

β_2-Glycoprotein I is a perchloric acid soluble protein in normal human plasma, first purified in 1961 (Schultze et al., 1961). Later, β_2-glycoprotein I has been purified from rat plasma (Polz et al., 1980). The complete amino acid sequence of human β_2-glycoprotein I has been determined by peptide sequencing (Lozier et al., 1984) as well as by cDNA sequencing (Steinkasserer et al., 1991), and it reveals that β_2-glycoprotein I is a single-chain polypeptide, consisting of 326 amino acid residues, with four or five asparagine-linked oligosaccharide groups. The calculated molecular mass of the polypeptide chain is 36 281 Da which is considerably less than the 54 200 Da estimated by SDS-PAGE (Lozier et al., 1984). Sedimentation equilibrium studies of human β_2-glycoprotein I indicate a molecular mass of 40 000-48 000 Da (Heimburger et al., 1964; Finlayson et al., 1967). Glycosylation could account for the discrepancy between the calculated and the estimated molecular mass values. The cDNA derived amino

[†] This work has been supported by The Danish Cancer Society (88-043; T.H.), The NOVO Foundation (1988-12-19 and 1989-12-18; T.H.), The Danish Science Research Council (11-8256; T.K.), The Carlsberg Foundation (88-0274/20; T.K.), The Danish Cancer Society (90-025; L.S.-J.), and The Center for Eucaryotic Gene Regulation, University of Aarhus.

[‡] The nucleotide sequence data reported will appear in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under Accession Number X60065.

* Address correspondence to this author at NOVO-Nordisk, NOVO Allé, DK-2880 Bagsvaerd, Denmark.

[§] Deceased.

0006-2960/92/0431-3611$03.00/0

acid sequence of rat β_2-glycoprotein I was recently published (Aoyama et al., 1989). This sequence consists of 297 amino acid residues, and there are five potential N-glycosylation sites.

The physiological role of β_2-glycoprotein I is still unknown. It is, however, well-known that β_2-glycoprotein I binds to negatively charged surfaces and molecules such as vesicles of negatively charged phospholipids, heparin, and DNA (Wurm, 1984; Schousboe, 1983, 1988; Polz et al., 1979; Krøll et al., 1976). It has been suggested that the binding of β_2-glycoprotein I to anionic surfaces is responsible for its inhibition of the contact activation of blood coagulation in vitro (Schousboe, 1985, 1988; Henry et al., 1988). β_2-Glycoprotein I has also been reported to bind to activated platelets, thereby repressing the platelet prothrombinase (Nimpf et al., 1986) and the adenylate cyclase activities (Schousboe, 1980). The affinity of β_2-glycoprotein I for platelets seems to depend on its ability to bind to the negatively charged phospholipids which appear on the outside of the membrane upon platelet activation.

β_2-Glycoprotein I is also known as apolipoprotein H as it is observed to comigrate with lipoproteins in sedimentation studies of normal human plasma (Polz et al., 1979; Lee et al., 1983), and it was recently found that anti-phospholipid autoantibodies recognize a complex antigen of β_2-glycoprotein I bound to phospholipid (McNeil et al., 1990; Galli et al., 1990). Bovine β_2-glycoprotein I isolated from fetal calf serum has recently been shown to inhibit the thymidine uptake in bovine erythroid cells (Li et al., 1990).

The amino acid sequence of β_2-glycoprotein I can be arranged into five mutually homologous domains, about 60 amino acid residues each. The consensus framework of the domains within β_2-glycoprotein I is based on the conservation of four half-cystine residues, two proline residues, two phenylalanine/tyrosine/histidine residues, two glycine residues, one leucine/valine residue, and one tryptophan residue, all in the same relative positions within each repeat. This type of domain, known as the Short Consensus Repeat (SCR), occurs frequently in a number of complement proteins and in a few noncomplement proteins as well [reviewed by Kristensen et al. (1987a) and Reid and Day (1989)].

Little information is available on the structural properties of SCR's in general. Lozier et al. (1984) have reported the location of 6 of the 11 disulfide bridges in human β_2-glycoprotein I. On the basis of these results, it has been suggested that within each SCR Cys-1 forms a disulfide bridge with Cys-3 and Cys-2 with Cys-4. This disulfide bridge pattern in SCR's was confirmed by Janatova et al. (1989) in the α-chains of human C4b-binding protein and by Hess et al. (1991) for human complement component C1s.

We present here the complete primary structure of bovine β_2-glycoprotein I determined by cDNA and peptide sequencing. We have determined the sites of N-linked glycosylation as well as the first complete disulfide bridge pattern of an SCR protein. In the first four SCR's from the N-terminus, the first and third and the second and fourth half-cystines are disulfide-linked, while in the fifth and most C-terminal SCR, the first and fourth, the second and fifth, and the third and sixth half-cystines are disulfide-linked.

Experimental Procedures

Materials for Protein Chemistry Work. Trypsin and chymotrypsin were from Boehringer Mannheim (Mannheim, Germany), *Staphylococcus aureus* V8 protease was from Worthington (Freehold, NJ), and thermolysin was from Sigma (St. Louis, MO). Heparin–Sepharose, DEAE-Sepharose CL-6B, and Sephadex G-50 SF were from Pharmacia (Uppsala, Sweden). Vydac C_{18} (10 μm and 5 μm) was from The Separation Group (Hesperia, CA); Nucleosil C_{18} (5 μm) was from Macherey-Nagel (Düren, Germany); Spherisorb ODS-2 (3 μm) was from Phase Separation (Deeside, U.K.); and Hypersil ODS (3 μm and 5 μm) was from Shandon (Cheshire, U.K.). Pth–C_{18} (5 μm) reversed-phase support, cartridge seals, and glass-fiber filters for the Applied Biosystems protein sequencers were from Applied Biosystems Inc. (Foster City, CA). Chemicals for amino acid sequence analysis were from Applied Biosystems or Rathburn (Walkerburn, Scotland, U.K.). Iodo[2-^{14}C]acetic acid was from Amersham International (Amersham, U.K.).

Materials for cDNA Cloning and Sequencing. Restriction enzymes and a random primed DNA labeling kit were from Boehringer Mannheim. A Sequenase kit was purchased from U.S. Biochemical Corp. (Cleveland, OH). Nitrocellulose filters (BA85) were from Schleicher & Schuell (Dassel, Germany), and radiolabeled nucleotides ([α-^{32}P]dATP and [α-^{35}S]dATPαS) were from Amersham International.

Purification of Bovine β_2-Glycoprotein I. Bovine plasma was obtained from fresh bovine blood collected in trisodium citrate (final concentration, 0.72% w/v). Perchloric acid was added to a final concentration of 0.15 M. After being stirred at 4 °C for 10 h, the β_2-glycoprotein I-containing supernatant was collected after centrifugation. Following extensive dialysis against 10 mM Tris-HCl, pH 8.2, the supernatant was further separated on a heparin–Sepharose column equilibrated in 10 mM Tris-HCl, pH 8.2. The heparin–Sepharose column was washed extensively in the equilibration buffer until A_{280} was below 0.01. The heparin binding proteins were eluted with a linear gradient of 0–0.4 M NaCl in 10 mM Tris-HCl, pH 8.2. The fractions containing β_2-glycoprotein I were identified by SDS–PAGE (Laemmli, 1970) followed by Western blotting (Towbin & Gordon, 1984) employing a rabbit anti-human β_2-glycoprotein I antibody kindly provided by Dr. Inger Schousboe, Panum Institute, University of Copenhagen, and a peroxidase-conjugated porcine anti-rabbit IgG. The β_2-glycoprotein I pool was dialyzed extensively against 10 mM Tris-HCl, pH 8.2, and further separated by anion-exchange chromatography on a DEAE-Sepharose column equilibrated with 10 mM Tris-HCl, pH 8.2, and eluted with a linear gradient of 0–0.35 M NaCl in 10 mM Tris-HCl, pH 8.2. The protein was judged to be pure by SDS–PAGE (Laemmli, 1970) and by N-terminal amino acid sequencing where a single sequence was observed. The amino acid sequence of the N-terminal 20 amino acid residues showed 95% identity with human β_2-glycoprotein I.

Proteolytic Cleavage of Bovine β_2-Glycoprotein I. Initially, parts of the amino acid sequence were determined from reduced and alkylated protein. Reduction and S-[^{14}C]-carboxymethylation of bovine β_2-glycoprotein I was performed according to Hirs (1967b). A 100-mg portion of S-[^{14}C]-carboxymethylated β_2-glycoprotein I was degraded with trypsin (1:50 w/w) in 0.1 M NH_4HCO_3, pH 8.2, overnight at 37 °C. The peptides were separated initially on a Sephadex G-50 SF column (5 cm × 120 cm) in 0.1 M NH_4HCO_3, pH 8.2. After being freeze-dried, all peptide pools were further separated by reversed-phase HPLC.

For the determination of the disulfide bridge pattern, two 20-mg portions of intact β_2-glycoprotein I were degraded with trypsin (1:50 w/w) in 0.1 M NH_4HCO_3, 2 M urea, pH 8.2, overnight at 37 °C. Although it is the general opinion that slightly alkaline conditions can result in disulfide exchanges, it is our experience during many years of research that this phenomenon rarely occurs in the absence of thio groups.

Titration of bovine β_2-glycoprotein I with 5,5′-dithiobis(2-nitrobenzoic acid) under denaturing conditions showed the absence of thio groups in the molecule. The peptides from one 20-mg portion were separated by Sephadex G-50 SF gel chromatography in 0.1 M NH_4HCO_3, pH 8.2 (recording at 230 nm not shown; Figure 2), while the peptides from the other 20-mg portion were directly separated by reversed-phase HPLC (Figure 2). Subdigests were carried out on isolated peptide fractions with the following enzymes and conditions: *S. aureus* V8 protease or chymotrypsin in 0.1 M NH_4HCO_3, pH 8.2, 4–24 h at 37 °C, or thermolysin in 100 mM pyridinium acetate, 1 mM $CaCl_2$, pH 6.5, 12 h at 65 °C.

Peptide separations by HPLC on reversed-phase columns (Vydac C_{18} or Nucleosil C_{18}; 4 × 250 mm) were carried out by employing a Hewlett-Packard 1084B liquid chromatograph eluting the material with linear gradients of increasing concentrations of ethanol, methanol, or acetonitrile in aqueous 0.1% TFA.

Amino Acid Composition Analysis. Following hydrolysis in 6 M HCl at 110 °C for 20 h in vacuo, amino acid compositions were determined using a Hewlett-Packard Amino Quant liquid chromatograph, equipped with an HP 1046 A programmable fluorescence detector. Cystine was determined as cysteic acid, obtained by performic acid oxidation (Hirs, 1967a) before hydrolysis in 6 M hydrochloric acid.

Amino Acid Sequence Analysis. Automated sequence analysis was performed using Applied Biosystems protein sequencers Models 470A and 477A according to the manufacturer's instructions.

Isolation and Sequence Analysis of a Bovine β_2-Glycoprotein I cDNA. An amplified cDNA library constructed from bovine liver poly(A^+) RNA in the pBR322 plasmid vector was kindly made available to us by Dr. Åke B. Lundwall, University of Lund, Sweden. Fifteen thousand colonies were screened in duplicate on blotted nitrocellulose filters (Hanahan & Meselson, 1980) with a human β_2-glycoprotein I cDNA probe (T. Kristensen et al., 1991) randomly labeled by using the Klenow enzyme, hexamer-oligodeoxynucleotides, and [α-^{32}P]dATP as label. The prehybridization and hybridizations were in a solution of 5× SSC, 10× Denhardt's solution, 10 μg/mL yeast RNA, and 0.5% sodium pyrophosphate at 65 °C under the same conditions as described earlier (Kristensen et al., 1987b). A clone (pBB2I) estimated to contain a full-length insert as judged by restriction analysis of several positives was chosen for sequencing. Selected restriction fragments covering the entire insert were subjected to sequencing by the dideoxynucleotide chain termination technique (Sanger et al., 1977) employing the Sequenase enzyme (Tabor & Richardson, 1987). The sequence results were aligned and compiled by using the DB programs of Staden (Staden, 1980, 1982a,b). The nucleotide and derived amino acid sequences were compared to the GenEMBL (release 26) and MIPSX (release 28) Databases, respectively (Devereux et al., 1984).

Results

β_2-Glycoprotein I cDNA Cloning. Fifteen thousand recombinant colonies of a bovine liver cDNA library constructed in the pBR322 plasmid vector were screened with a cDNA probe covering most of the coding sequence of human β_2-glycoprotein I (T. Kristensen et al., 1991). Plasmid DNA from eight hybridizing clones were examined by restriction analysis, and the clone (pBB2I) estimated to contain the largest insert was sequenced. The clone was not full-length (Figure 1), but it contained enough sequence to cover the 326 amino acid residues of mature β_2-glycoprotein I necessary to establish the primary structure. All together the insert encompasses (1) 48 bp encoding 16 amino acids of a hydrophobic signal peptide, lacking three amino acids when compared to the recently published rat (Aoyama et al., 1989) and human (Steinkasserer et al., 1991) β_2-glycoprotein I cDNA-derived amino acid sequence; (2) 978 bp encoding the mature protein; and (3) following the stop codon TAA, 100 bp of 3′-nontranslated sequence including the canonical polyadenylation signal sequence AATAAA located 29 bp upstream from the poly(A) tail. Two potential polyadenylation signal sequences, AATTAA and ATTAAA, were seen starting at positions 1048 and 1049, respectively.

Table I: Amino Acid Sequences of Peptides Obtained from Reduced and S-[^{14}C]Carboxymethylated β_2-Glycoprotein I

amino acid sequence	position in β_2-glycoprotein I
GRTCPKPDELPFSTVVPLKR[a]	1–20
TYEPGEQIVFSCQPGYVSR[b]	21–39
RFTCPLTGLWPINTLK[b]	44–59
♦	
CMPRVCPFAGILENGTVRYTTF[b]	60–81
CTEEGKWSPDLPVCAPITCP[b]	105–124
♦	
FASLSVYKPLAGNN[b]	131–144
♦ ♦	
AMFGNDT(E/V)VTCTEHGNWTQLPECR[b,c]	160–182
CPFPSRPDNGFVN[b]	186–198
DTATFGCHETYSLDGPEEVECSK[b]	209–231
ATVIYEGER[b]	252–260
CSYTEDAQCIDGTIEIPK[b]	288–305

[a] N-Terminal sequence of S-[^{14}C]carboxymethylated β_2-glycoprotein I. [b] Tryptic peptide from S-[^{14}C]carboxymethylated β_2-glycoprotein I. [c] Equal amounts of Pth-Glu and Pth-Val were observed in position 8 of this peptide. ♦, N-bound carbohydrate.

Amino Acid Sequence of Bovine β_2-Glycoprotein I. In order to confirm the amino acid sequence deduced from the cDNA sequence and in order to determine the sites of N-glycosylation and the locations of disulfide bridges, we have sequenced several peptides all together covering 242 of the 326 amino acid residues of bovine β_2-glycoprotein I (Tables I and II; shown in bold italics in Figure 1).

From the tryptic digest of reduced and S-[^{14}C]carboxymethylated protein, eight pools were collected after gel filtration chromatography on a column of Sephadex G-50 SF. Each of these pools was subjected to further separation by reversed-phase HPLC, and a total of eleven peptides were isolated and sequenced (Table I). At position 167, Pth-Val and Pth-Glu (Table I and Figure 1) appeared in equal amounts, indicating heterogeneity at this position.

N-Linked Carbohydrate Attachments. Peptides containing Asn-linked oligosaccharides were identified by observing *N*-acetylglucosamine in peptide hydrolysates. Sequence analysis of such peptides showed that Asn-linked oligosaccharides are attached to the Asn residues in positions 73, 143, 164, 174, and 234 (Figure 1). All attachment sites are located in the carbohydrate acceptor consensus sequence Asn-Xaa-Ser/Thr. In the cases of Asn174 and Asn234, the second position of the consensus sequence was occupied by a Trp residue, which rarely occurs at this position (Gavel & von Heijne, 1990). Four of the attachment sites (Asn143, Asn164, Asn174, and Asn234) are identical to those found in both rat and human β_2-glycoprotein I (Aoyama et al., 1989; Lozier et al., 1984; Steinkasserer et al., 1991).

Disulfide Bridge Pattern of Bovine β_2-Glycoprotein I. The sequences of isolated disulfide-bonded peptides necessary to assign the 11 disulfide bridges of β_2-glycoprotein I are summarized in Table II.

The amino acid sequences of disulfide-bonded peptides were determined by simultaneous Edman degradation of the two

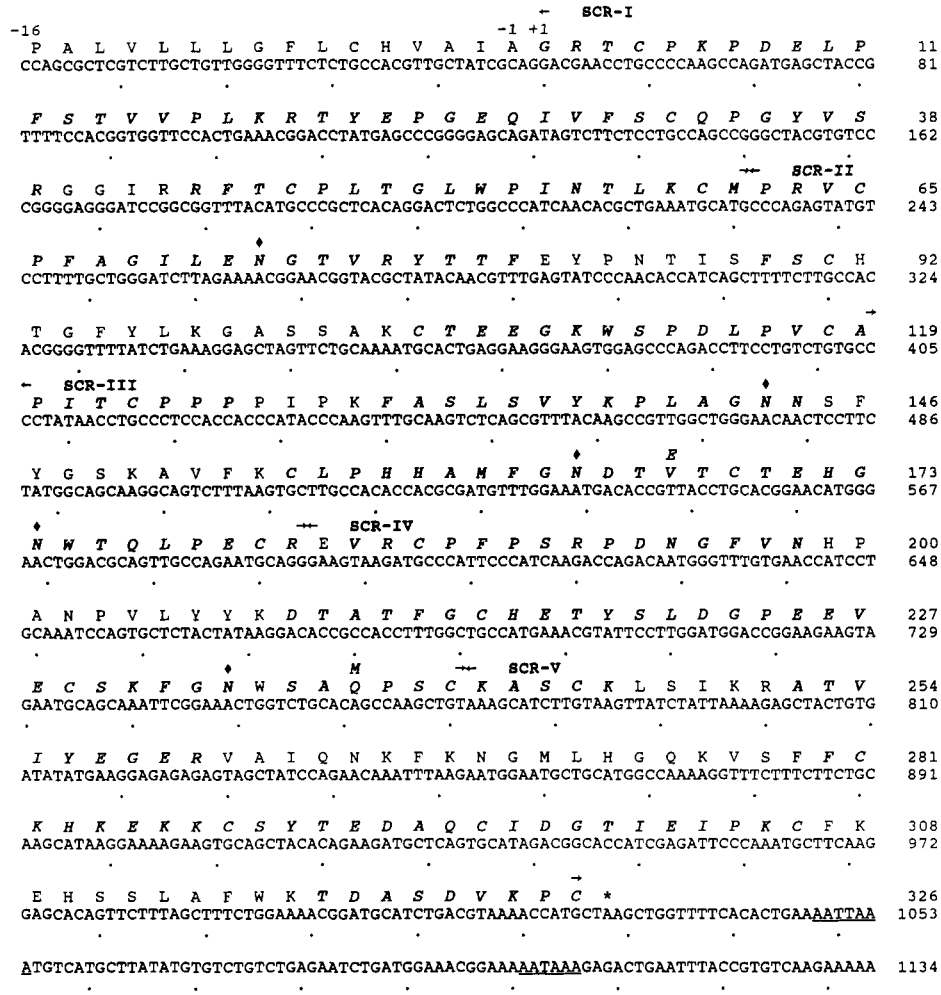

FIGURE 1: The nucleotide (cDNA) and derived amino acid sequences of bovine β_2-glycoprotein I. Amino acid sequences also determined by sequencing selected peptides (see Tables I and II) derived from proteolytically cleaved β_2-glycoprotein I are shown in bold italics. Sites of N-glycosylation are indicated by diamonds (♦). The stop codon, TAA, is indicated with an asterisk, and the polyadenylation signal sequence, AATAAA, is underlined. Two overlapping potential polyadenylation signal sequences are underlined. The span of each Short Consensus Repeat is indicated by arrows.

linked peptide chains. The position of the first half-cystine in a disulfide bridge is observed as a missing Pth derivative in the actual sequencing step, while the second half-cystine is released as Pth-cystine. This derivative is observed as a sharp peak eluting in front of Pth-Tyr in the standard HPLC separation of Pth-amino acids.

The disulfide bridges are designated I–XI in order of appearance of the first Cys residue in each disulfide bridge. In Figure 2 are outlined the routes of purification of disulfide-bridged peptides covering all 11 disulfide bridges. One portion of β_2-glycoprotein I (Figure 2) was degraded with trypsin as described in Experimental Procedures, and the resulting peptides were separated by gel filtration (not shown) into five fractions. Fractions 1–4 (numbered according to appearance from the gel filtration column) containing substantial amounts of cystine were further separated by HPLC. The peptides containing disulfide bridges I (derived from fraction 3), II (from fraction 4), III (from fraction 1), and VIII (from fraction 2) were purified and directly sequenced (Table II). Peptide material in a major cystine-containing peak from the HPLC purification was further degraded with *S. aureus* V8 protease, and the peptides were separated by HPLC. Peptides containing disulfide bridges IX and X were identified (Table II). Another portion of β_2-glycoprotein I was also treated with trypsin (Figure 2), and the resulting peptides were directly fractionated by HPLC (not shown). A part of the material in one major, large cystine-containing peak was degraded with chymotrypsin before peptide separation by HPLC. The peptide containing disulfide bridge VI was isolated and sequenced (Table II). The rest of the material from this peak was treated with thermolysin, and the peptides were fractionated by HPLC. Peptides containing disulfide bridges IV, V, VII, and XI, respectively, were isolated and sequenced (Table II).

Primary Structure of Bovine β_2-Glycoprotein I. Figure 1 shows the complete amino acid and nucleotide sequences of bovine β_2-glycoprotein I. The results of sequence analysis of peptides isolated from enzymatic digests of S-[^{14}C]carboxymethylated or intact bovine β_2-glycoprotein I are shown in bold italic letters. Only one discrepancy was found between the amino acid sequence deduced from the cDNA sequence and the amino acid sequence determined at the protein level. At position 238, the nucleotide sequence predicts a glutamine while a methionine was determined in the tryptic peptide containing disulfide bridge VIII (Table II and Figure 1). The N-terminal 22 amino acid residues sequenced by Li et al. (1990) match with the sequence presented here.

The most notable differences between the bovine (and rat) and human amino acid sequences were found at positions 102 and 169, where a half-cystine and a glycosylated asparagine, respectively, were determined in the human amino acid sequence (Lozier et al., 1984). In bovine β_2-glycoprotein I, positions 102 and 169 are occupied by a serine and a half-cystine, respectively. It should be noted that the cDNA-de-

Table II: Sequences of Disulfide Bridged Peptides from β_2-Glycoprotein I

amino acid sequences	positions of half-cystines	disulfide bridge no.	enzymes used for digest
TCPKPDELPFS FTCPLTGLWPI	4; 47	I	tr
TYEPGEQIVFSCQPGYVSR CMPR ♦	32; 60	II	tr
VCPFAGILENGTVR CTEEGK	65; 105	III	tr
FSC VCAP	91; 118	IV	tr + th
ITCPPP (E/V)TCTEHG[a]	123; 169	V	tr + th
CLPHHAMFG TQLPEC	155; 181	VI	tr + ct
VRCPFPSRPDNG VECSK	186; 229	VII	tr + th
DTATFGCHET FGNXSAMPSCK ♦	215; 241	VIII	tr
ASCK DAQCIDGTI	245; 296	IX	tr + sp
FCKHKEKK IEIPKC	281; 306	X	tr + sp
CSYTE TDASDVKPC	288; 326	XI	tr + th

[a] Equal amounts of Pth-Val and Pth-Glu were observed for position 167. tr, trypsin; ct, chymotrypsin; th, thermolysin; sp, *Staphylococcus aureus* V8 protease; ♦, N-bound carbohydrate.

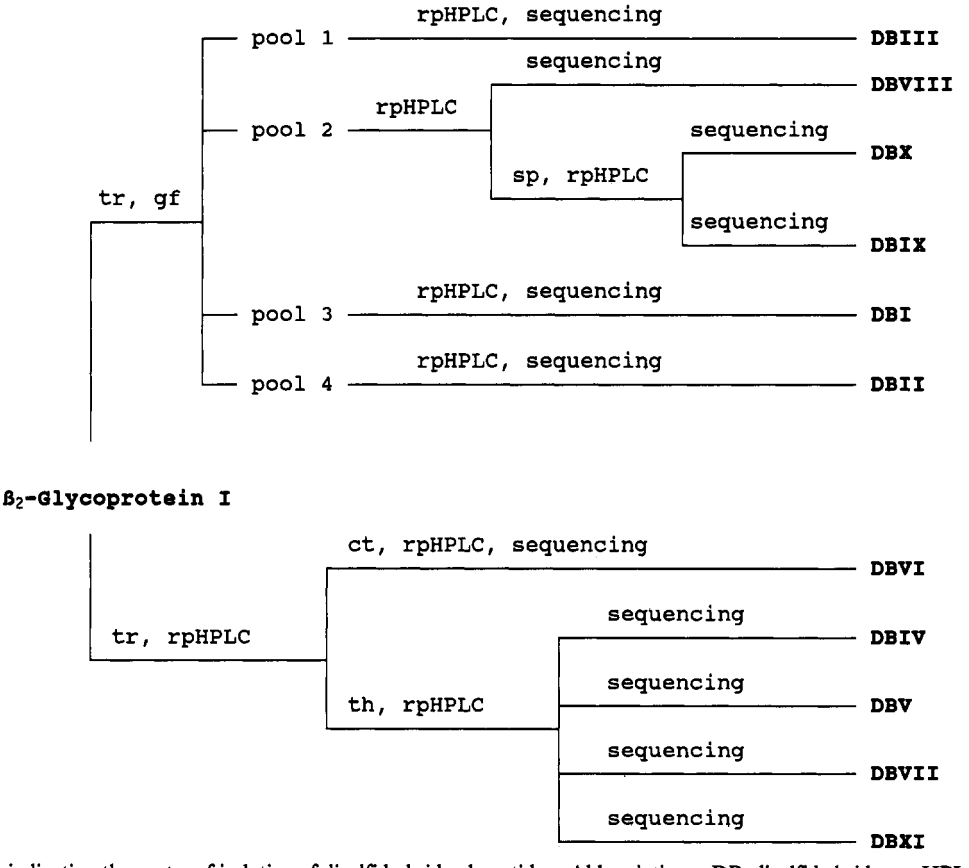

FIGURE 2: Flow scheme indicating the routes of isolation of disulfide-bridged peptides. Abbreviations: DB, disulfide bridge; rpHPLC, reversed-phase HPLC; gf, gel filtration; tr, trypsin; ct, chymotrypsin; sp. *S. aureus* V8 protease; th, thermolysin.

rived amino acid sequences of rat and human β_2-glycoprotein I (Aoyama et al., 1989; Steinkasserer et al., 1991; T. Kristensen et al., 1991) also predict a serine at position 102 and a cysteine at position 169. Another strong argument in favor

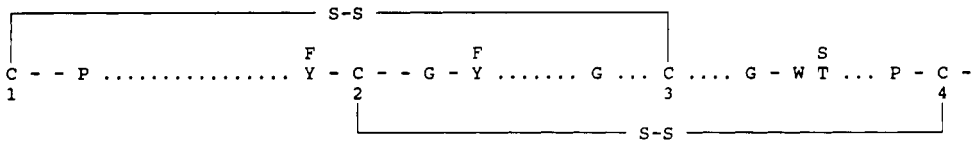

FIGURE 3: The general consensus sequence of the Short Consensus Repeat. The intradomain disulfide bridges connecting Cys-1 and Cys-3 and Cys-2 and Cys-4 are shown. The hyphen indicates a space corresponding to one amino acid residue.

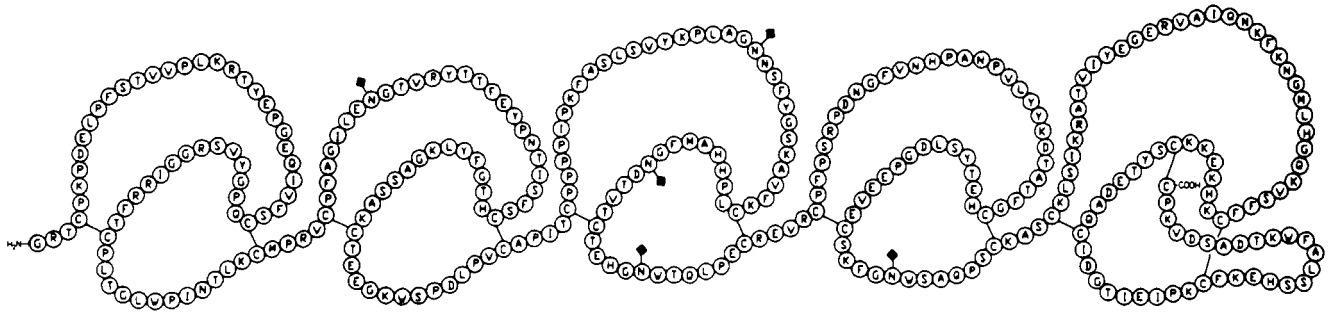

FIGURE 4: A schematic representation of bovine β_2-glycoprotein I showing the organization of the molecule into five Short Consensus Repeats each with two intradomain disulfide bridges except for the C-terminal domain which contains three intradomain disulfide bridges. The N-linked glycosylation positions are indicated with diamonds (♦).

of a serine at position 102 and a half-cystine at position 169 is that the occurrences of these residues at these positions are consistent with the SCR consensus of four invariant half-cystines in each SCR (Figure 3).

Discussion

We have identified the 11 disulfide bridges in bovine β_2-glycoprotein I. The amino acid sequence of β_2-glycoprotein I can be divided into five mutually homologous SCR domains, and this report confirms the suggestion that SCR domains contain only intradomain disulfide bridges. The four half-cystines within each of the N-terminal four SCR domains form disulfide bridges in a 1–3, 2–4 pattern. In the last and fifth SCR domain, three disulfide bridges are present, but the disulfide bridge pattern in the fifth domain is still consistent with the 1–3, 2–4 pattern. The two additional half-cystines (Cys288 and Cys326) forming disulfide bridge XI are not part of the SCR consensus sequence (Figure 3).

In 1984, the complete amino acid sequence of human β_2-glycoprotein I was published, and 6 of the 11 disulfide bridges were identified (Lozier et al., 1984). Five of the disulfide bridges identified in human β_2-glycoprotein I correspond to the disulfide bridges I, II, III, VI, and VII in bovine β_2-glycoprotein I, while the sixth disulfide bridge connecting Cys281 and Cys288 is inconsistent with our results. In bovine β_2-glycoprotein I, we find that Cys281 is forming a disulfide bond with Cys306 (disulfide bridge X), while Cys288 is connected to the C-terminal Cys326 (disulfide bridge XI). These results may indicate a difference in the disulfide bridge pattern of the fifth SCR between these two species.

Recently, the positions of several of the disulfide bridges in the eight SCR domains of the human C4b-binding protein α-chains (Janatova et al., 1989) as well as those of all of the disulfide bridges of human complement C1s (Hess et al., 1991) were determined. The results that Cys-1 is disulfide-bound to Cys-3 and Cys-2 to Cys-4 within each SCR are consistent with the results obtained for bovine and human β_2-glycoprotein I.

Our results and the results of Lozier et al. (1984), Janatova et al. (1989), and Hess et al. (1991) together with an increasing knowledge about gene structures of SCR proteins [e.g., complement factor H, C4b-binding protein, complement receptors type 1 and 2, interleukin-2 receptor (Barnum et al., 1989; Fujisaku et al., 1989; Leonard et al., 1985; Vik et al., 1988; Wong et al., 1989)] strongly suggest that the SCR is a separate physical and genetical entity.

In summary, we present here the first complete disulfide bridge pattern of a protein composed entirely of SCR domains. The results show that the amino acid sequence of β_2-glycoprotein I is arranged into five mutually homologous domains with two disulfide bridges in each of the first four domains (Cys-1 to Cys-3 and Cys-2 to Cys-4), and three disulfide bridges in the last domain (Cys-1 to Cys-4, Cys-2 to Cys-5, and Cys-3 to Cys-6) as shown in the schematic representation of bovine β_2-glycoprotein I in Figure 4.

Added in Proof

After submission of this paper, Kato and Enjyoji (1991) published the amino acid sequence and location of the disulfide bonds in bovine β_2-glycoprotein I. The results of these authors are in agreement with the results presented here except for the assignment of amino acid residue 283. At this position we find a His residue, whereas Kato and Enjyoji find an Asn residue.

Acknowledgments

We thank Rita Rosendahl Hansen for excellent technical assistance.

References

Aoyama, A., Chan, Y.-L., & Wool, I. G. (1989) *Nucleic Acids Res. 17*, 6401.

Barnum, S. R., Kristensen, T., Chaplin, D. D., Seldin, M. F., & Tack, B. F. (1989) *Biochemistry 28*, 8312–8317.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Finlayson, J. S., & Mushkinski, J. F. (1967) *Biochim. Biophys. Acta 147*, 413–420.

Fujisaku, A., Harley, J. B., Frank, M. B., Gruner, B. A., Frazier, B., & Holers, V. M. (1989) *J. Biol. Chem. 264*, 2118–2125.

Galli, M., Comfurius, P., Maassen, C., Hemker, H. C., de Baets, M. H., van Breda Vriesman, P. J., Barbui, T., Zwaal, R. F., & Bevers, E. M. (1990) *Lancet 335*, 1544–1547.

Gavel, Y., & von Heijne, G. (1990) *Protein Eng. 3*, 433–442.

Hanahan, D., & Meselson, M. (1980) *Gene 10*, 63–67.

Heimburger, N., Heide, K., Haupt, H., & Schultze, H. E. (1964) *Clin. Chim. Acta 10*, 293–307.

Henry, M. L., Everson, B., & Ratnoff, O. D. (1988) *J. Lab. Clin. Med. 111*, 519–523.
Hess, D., Schaller, J., & Rickly, E. E. (1991) *Biochemistry 30*, 2827–2833.
Hirs, C. H. W. (1967a) *Methods Enzymol. 11*, 197–199.
Hirs, C. H. W. (1967b) *Methods Enzymol. 11*, 199–202.
Janatova, J., Reid, K. B. M., & Willis, A. C. (1989) *Biochemistry 28*, 4754–4761.
Kato, H., & Enjyoji, K. (1991) *Biochemistry 30*, 11687–11694.
Kristensen, T., D'Eustachio, P., Ogata, R. T., Chung, L. P., Reid, K. B. M., & Tack, B. F. (1987a) *Fed. Proc. 46*, 2463–2469.
Kristensen, T., Ogata, R. T., Chung, L. P., Reid, K. B. M., & Tack, B. F. (1987b) *Biochemistry 26*, 4668–4674.
Kristensen, T., Schousboe, I., Boel, E., Mulvihill, E. M., Hansen, R. R., Møller, K. B., Møller, N. P. H., & Sottrup-Jensen, L. (1991) *FEBS Lett. 289*, 183–186.
Krøll, J., Larsen, J. K., Loft, H., Ezban, M., Wallevik, K., & Faber, M. (1976) *Biochim. Biophys. Acta 434*, 490–501.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lee, N. S., Brewer, B. H., Jr., & Osborne, J. C., Jr. (1983) *J. Biol. Chem. 258*, 4765–4770.
Leonard, W. J., Depper, J. M., Kanehisa, M., Krönke, M., Peffer, N. J., Svetlik, P. B., Sullivan, M., & Green, W. C. (1985) *Science 230*, 633–639.
Li, Q., Blacher, R., Esch, F., & Congote, L. F. (1990) *Biochem. J. 267*, 261–264.
Lozier, J., Takahashi, N., & Putnam, F. W., (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3640–3644.
McNeil, H. P., Simpson, R. J., Chesterman, C. N., & Krilis, S. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4120–4124.
Nimpf, J., Bevers, E. M., Bomans, P. H. H., Till, U., Wurm, H., Kostner, G. M., & Zwaal, R. F. A. (1986) *Biochim. Biophys. Acta 884*, 142–149.
Polz, E., & Kostner, G. M. (1979) *Biochem. Biophys. Res. Commun. 90*, 1305–1312.
Polz, E., Wurm, H., & Kostner, G. M. (1980) *Int. J. Biochem. 11*, 265–270.
Reid, K. B. M., & Day, A. J. (1989) *Immunol. Today 10*, 177–180.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schousboe, I. (1980) *Thromb. Res. 19*, 225–237.
Schousboe, I. (1983) *Int. J. Biochem. 15*, 1393–1401.
Schousboe, I. (1985) *Blood 66*, 1086–1091.
Schousboe, I. (1988) *Int. J. Biochem. 20*, 309–315.
Schultze, H. E., Heide, K., & Haupt, H. (1961) *Naturwissenschaften 48*, 719.
Staden, R. (1980) *Nucleic Acids Res. 8*, 3673–3694.
Staden, R. (1982a) *Nucleic Acids Res. 10*, 2951-2961.
Staden, R. (1982b) *Nucleic Acids Res. 10*, 4731–4751.
Steinkasserer, A., Estaller, C., Weiss, E. H., Sim, R. B., & Day, A. J. (1991) *Biochem. J. 277*, 387–391.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.
Towbin, H., & Gordon, J. (1984) *J. Immunol. Methods 72*, 313–340.
Vik, D. P., Keeney, J. B., Munos-Canoves, P., Chaplin, D. D., & Tack, B. F. (1988) *J. Biol. Chem. 263*, 16720–16724.
Wong, W. W., Cahill, J. M., Rosen, M. D., Kennedy, C. A., Bonaccio, C. B., Morris, M. J., Wilson, J. G., Klickstein, L. B., & Fearon, D. T. (1989) *J. Exp. Med. 169*, 847–863.
Wurm, H. (1984) *Int. J. Biochem. 16*, 511–515.

Stable Substructures of Eightfold $\beta\alpha$-Barrel Proteins: Fragment Complementation of Phosphoribosylanthranilate Isomerase†

Jörg Eder and Kasper Kirschner*
Abteilung für Biophysikalische Chemie, Universität Basel, CH-4056 Basel, Switzerland
Received September 3, 1991; Revised Manuscript Received January 10, 1992

ABSTRACT: The $(\beta\alpha)_8$ (or "TIM")-barrel protein phosphoribosylanthranilate isomerase from *Saccharomyces cerevisiae* was cleaved between the sixth and seventh $\beta\alpha$ module to test the capacity of the resulting fragments to adopt native format autonomously. The fragments, which were expressed from separate coding sequences, were soluble and monomeric. The amino-terminal fragment p_1 was compact, possessed an almost nativelike far-UV but a strongly reduced near-UV CD spectrum, and unfolded cooperatively with guanidinium chloride. In contrast, the carboxyl-terminal fragment p_2 was less compact than fragment p_1, possessed only a weak far-UV and no detectable near-UV CD spectrum, and unfolded noncooperatively. The fragments assembled stoichiometrically to a complex with $K_d = 0.2$ μM, which was enzymically almost fully active. The rate of assembly was limited by a first-order process, probably the isomerization of the carboxyl-terminal fragment p_2 to an assembly-competent structure. These results support a folding mechanism that comprises an intermediate with the first six $\beta\alpha$ units folded in roughly native format and the last two $\beta\alpha$ units partially unfolded. The similar behavior of the analogous fragments of the α subunit of tryptophan synthase supports the hypothesis that these two $(\beta\alpha)_8$-barrel proteins have evolved from a common ancestor.

Denatured proteins refold in vitro by collapsing to a compact state that subsequently rearranges itself to the native structure. This "molten globule" (Kuwajima, 1989) or "collapsed intermediate" (Kim & Baldwin, 1990) state can possess a circular dichroism spectrum in the peptide absorbance region that suggests a content of secondary structure similar to that of the stable folded state. Nativelike secondary and supersecondary structures of segments of the polypeptide chain could

† This work was supported in part by Grant 31.25711.88 from the Swiss National Science Foundation.
* To whom correspondence should be addressed.

0006-2960/92/0431-3617$03.00/0