Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 663–667, January 1992
Biochemistry

Amino acid sequence of human plasma galactoglycoprotein: Identity with the extracellular region of CD43 (sialophorin)

(glycans)

KARL SCHMID[a], MATTHIAS A. HEDIGER[b], REINHARD BROSSMER[c], JOHN H. COLLINS[d], HEINZ HAUPT[e], THOMAS MARTI[a,f], GWYNNETH D. OFFNER[a], JOHANN SCHALLER[g], KEIICHI TAKAGAKI[a,h], MARY T. WALSH[i], H. GERHARD SCHWICK[e], FRED S. ROSEN[j], AND EILEEN REMOLD-O'DONNELL[j, k]

[a]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118; [b]Brigham and Women's Hospital, Renal Division, Harvard Medical School, Boston, MA 02115; [c]Department of Biochemistry, Faculty of Medicine, University of Heidelberg, 6900 Heidelberg, Federal Republic of Germany; [d]Department of Biological Chemistry, School of Medicine, and Medical Biotechnology Center, Maryland Biotechnology Institute, University of Maryland, Baltimore, MD 21201; [e]Behringwerke AG, Marburg/Lahn, Federal Republic of Germany; [g]Institut für Biochemie, University of Bern, CH-3012 Bern, Switzerland; [h]Department of Biophysics, Boston University School of Medicine, Boston, MA 02118; and [j]The Center for Blood Research and Departments of Biological Chemistry and Molecular Pharmacology and Pediatrics, Harvard Medical School, Boston, MA 02115

Communicated by Emil R. Unanue, October 18, 1991

ABSTRACT **The amino acid sequence of galactoglycoprotein purified from human plasma was elucidated to 75% completeness by using chemical degradation of peptides and glycopeptides derived from digests of the protein with seven specific proteases. This sequence represents a polypeptide chain of approximately 220 amino acid residues including a high content of serine, threonine, alanine, and proline with one N-linked and multiple O-linked glycans. Comparison of peptide sequences from the native galactoglycoprotein and the deglycosylated derivative demonstrated the locations of 25 sites of O-glycosylation and three serine sites that are not glycosylated. The homogeneous N terminus was established as serine. C-terminal analysis revealed multiple C-terminal residues, suggesting that galactoglycoprotein molecules are of varying lengths. A search of a protein data base revealed that the galactoglycoprotein polypeptide is identical to the N-terminal (extracellular) polypeptide region of the blood-cell surface molecule CD43 (sialophorin, leukosialin). Further support of the relatedness of these molecules was obtained by immunoprecipitation of ^{125}I-labeled galactoglycoprotein by monoclonal anti-CD43 antibodies. The composition and properties of the molecules together with the known structure of the gene encoding CD43 suggest that galactoprotein is derived by proteolytic cleavage from transmembrane "hexasaccharide CD43," known to be expressed on neutrophils, activated T lymphocytes, and platelets.**

In 1980 we described the isolation from normal human blood plasma of galactoglycoprotein (Galgp) and its chemical and physicochemical properties (1). This molecule possesses an unusual carbohydrate content of 76% consisting mainly of *N*-acetylneuraminic acid (NeuAc), galactose (Gal), *N*-acetylgalactosamine (GalNAc), and *N*-acetylglucosamine (GlcNAc) (1) and is comprised essentially of O-glycosidic hexasaccharides (2). The single polypeptide chain of approximately 220 residues is composed primarily of threonine, serine, proline, glycine, and alanine. Because of its high Gal and GalNAc contents, Galgp was considered to be atypical of plasma proteins as these are generally devoid of GalNAc and have a relatively high mannose content (1, 2). Hence, it was speculated that Galgp, whose blood plasma concentration is greater than 10 μg per ml (3), might be derived from the surface of blood cells (1).

Recently, a glycoprotein with carbohydrate content greater than 60% was isolated from rodent (4) and human (5–7) leukocytic cells and was designated CD43 (sialophorin, leukosialin). CD43 is a major blood-cell surface molecule of lymphocytes, monocytes, and neutrophils (8) and is associated with important *in vitro* functions including cytoplasmic signaling (9), proliferation of T lymphocytes (10), and activation of monocytes (11) and natural killer cells (12) (CD43 is reviewed in ref. 13). Defects of CD43 have been detected in the inherited immunodeficiency Wiskott–Aldrich syndrome (14, 15), and anti-CD43 antibodies are commonly present in patients infected with human immunodeficiency virus type 1 (HIV-1) (16). There is considerable evidence for the expression on human blood cells of two CD43 species with different O-glycosylation of a common polypeptide (6, 8, 17). CD43 molecules with O-glycans that are predominantly tetrasaccharides (tetrasaccharide CD43) are found on circulating T lymphocytes and monocytes, and CD43 molecules with O-linked glycans that are predominantly hexasaccharides (hexasaccharide CD43) are found on neutrophils, activated T lymphocytes, and at low levels on platelets (8, 18).

More recently, the amino acid sequences of human and rat CD43 were elucidated by using mainly cloning techniques (19–21). The human sequence defined a mature transmembrane protein consisting of three regions: (*i*) a 235-amino acid extracellular N-terminal region that is highly O-glycosylated and possesses repeating units, (*ii*) a 23-residue hydrophobic transmembrane region, and (*iii*) a 123-residue cytoplasmic region (19, 20).

We describe in this manuscript the preparation of peptides and glycopeptides from human plasma Galgp and the determination of their amino acid sequences. Based on amino acid sequence and certain molecular characteristics, the Galgp polypeptide is shown to be identical to the N-terminal extracellular polypeptide region of CD43.

MATERIALS AND METHODS

Materials. Galgp was isolated from normal human plasma (1) with the additional step of solid-phase immunoabsorption to remove small amounts of proteinaceous contaminants. The antiserum was raised in a rabbit that had been immunized with purified Galgp but that produced antibodies only against

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Galgp, galactoglycoprotein; NeuAc, *N*-acetylneuraminic acid; GalNAc, *N*-acetylgalactosamine; mAb, monoclonal antibody.

[f]Present address: Departments of Biology and Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139.
[h]Present address: Department of Biochemistry, Hirosaki University, School of Medicine, Hirosaki 036, Japan.
[k]To whom reprint requests should be addressed.

the "impurities." Asialo and a-*O*-glyco derivatives of the protein were prepared enzymatically as described (22). L10 is a mouse monoclonal IgG1 reactive with a glycosylation-dependent, sialic acid-independent epitope in the N-terminal region of CD43 (15, 23). L60 (anti-Leu-22) (Becton Dickinson) is a mouse monoclonal IgG1 reactive with a sialic acid-dependent epitope of CD43 (24).

Methods. The strategy for amino acid sequence determination has been presented (25, 26). Peptides and glycopeptides were prepared enzymatically from Galgp by using chymotrypsin, lysyl endopeptidase, porcine pancreatic elastase, prolyl endopeptidase, *Staphylococcus aureus* V8 protease, thermolysin, and trypsin. Galgp and its derivatives were also cleaved with CNBr. The initial separation of the peptides and glycopeptides was achieved by gel filtration followed by ion-exchange chromatography on DEAE-cellulose, and the final purification was done by HPLC with various solvents and columns. Automated Edman degradation was performed with a Beckman model 890C sequencer and an Applied Biosystem model 470A sequencer or as described (27). The primary amino acid sequence of Galgp was analyzed for elements of secondary structure by using GCT sequence analysis software (version 6.1, August 19, 1989, copyright John Devereux) (28).

RESULTS AND DISCUSSION

Amino Acid Sequence. The amino acid sequence of seven regions of Galgp totaling 171 residues was elucidated by sequencing peptides and glycopeptides isolated from the proteolytic digests and from CNBr cleavage of Galgp (Fig. 1). The identification of the N-terminal residue presented difficulties when analyzing the native glycoprotein. However, by using the asialo derivative, this residue was established unambiguously as serine. The N-terminal 72 residues were established by using overlapping peptides as follows. Asialo-Galgp afforded the N-terminal eight residues with unidentified residues at positions 2, 3, and 7, which were shown to be glycosylated threonine residues by analyzing the derivative from which the O-glycans were enzymatically removed. A fragment generated by *S. aureus* V8 protease afforded residues 1–36. Additional peptides supported this sequence. A CNBr fragment starting at residue 26 extended the sequence to residue 43, and two further peptides afforded extensions to residues 54 (peptide Ch-2-1) and 58 (peptide T-112). The *S. aureus* V8 peptide V-p2 extended the sequence from residue 55 to residue 72. Although digests obtained by seven proteases were analyzed, no overlapping peptide was recovered to extend the sequence beyond residue 72. The next segment that was defined with overlapping peptides reached from residue 74 to residue 104. Additional fragments generated by thermolysin (Th-c8 and Th-d), cyanogen bromide (CB-11), chymotrypsin (Ch-8), and elastase (E-d) provided the sequences of five additional regions (Fig. 1).

When compared with sequences in a protein sequence data bank, the Galgp sequences were found to be identical to sequences of the blood-cell surface molecule CD43 (19, 20). By the insertion of spaces between the sequenced regions as shown in Fig. 1, the Galgp polypeptide can be seen to be colinear with the N-terminal (extracellular) polypeptide region of CD43 (Fig. 2).

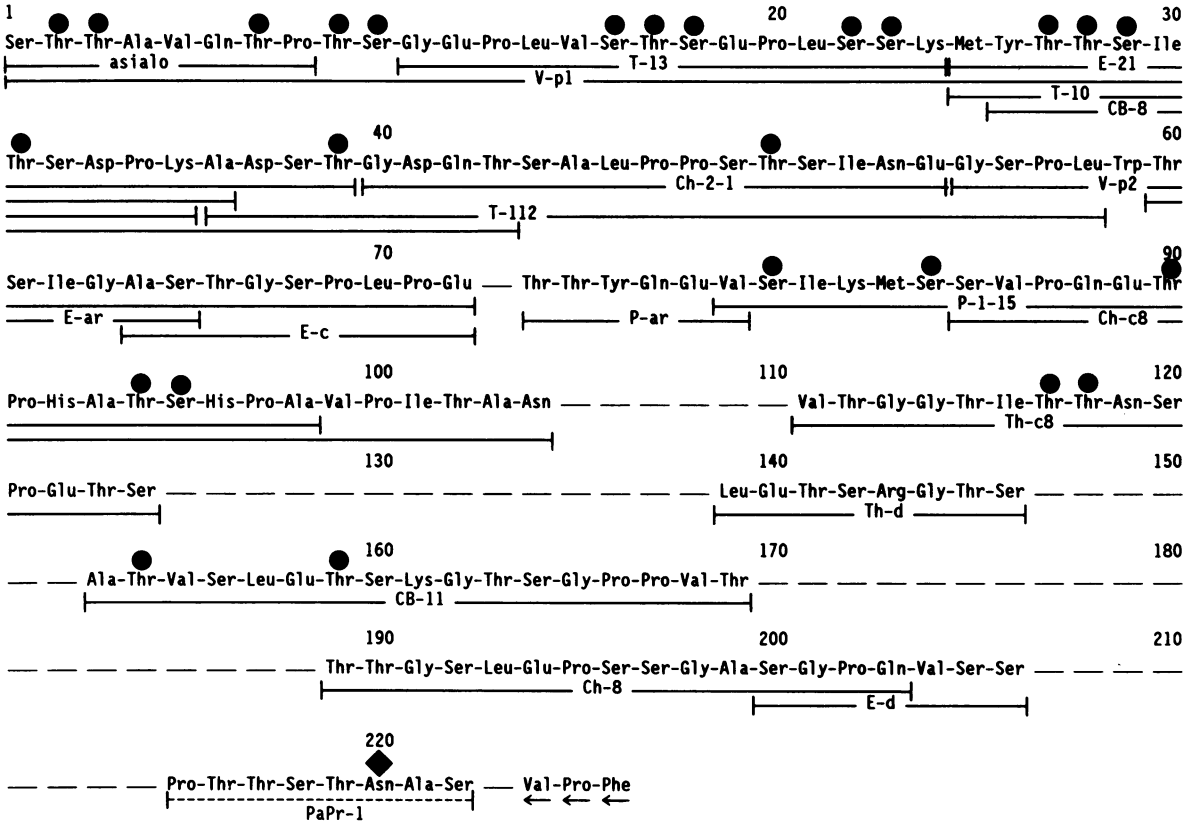

FIG. 1. Amino acid sequence of human plasma Galgp. Solid lines represent sequenced peptides generated by cyanogen bromide (CB), chymotrypsin (Ch), elastase (E), prolyl endopeptidase (P), thermolysin (Th), trypsin (T), and V8 protease from *S. aureus* (V). For peptide PaPr-1, which was isolated from a papain–pronase digest (2), only the composition was determined. The basis for the proposed sequence of PaPr-1 and the insertion of spaces between sequenced regions are explained in the text. The arrows pointing to the left indicate amino acids released by carboxypeptidase Y. The filled circles indicate serine and threonine residues found to carry O-glycans; the serine residues at positions 32, 49, and 51 were found to be nonglycosylated. The glycosylation state of the remaining serine and threonine residues was not resolved. The diamond at Asn-220 indicates an N-glycan. Additional sequences confirmed the positions of most of the amino acid residues.

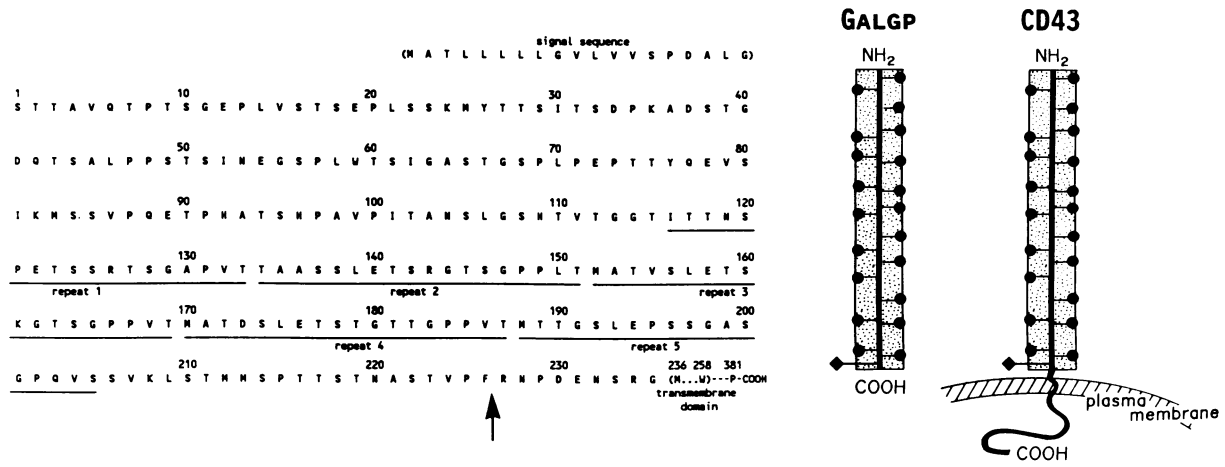

FIG. 2. Comparison of the amino acid sequences of Galgp and CD43. (*Left*) Extracellular 235 amino acids of CD43 (381 total amino acids) (19, 20). By spacing the sequenced peptides as indicated in Fig. 1, each of the 171 amino acids that were determined for Galgp is identical to the corresponding residue of extracellular CD43. The region of five tandem repeats of 18 amino acids, Ile-116 to Ser-205 (19), is underlined. The C terminus corresponding to the best characterized Galgp species (CD43 residues 1–226) (described in the text) is indicated by an arrow at Phe-226. (*Right*) Models of Galgp and CD43. Circles at arbitrary positions represent the O-glycans, and the diamond represents the N-glycan at Asn-220. The region of CD43, which is O-glycosylated (residues 1–223), and all of Galgp are shown as expanded structures as predicted for mucin-like molecules (29, 30) and as demonstrated recently for Galgp (ref. 22 and text) and CD43 (31).

As previously noted (19), the CD43 region between residues 116 and 205 consists of five repeats of an 18-amino acid unit (Fig. 2, underlined sequences). A Galgp peptide was sequenced (residues 111–124) that includes nine residues of repeat 1. Of repeat 2, an eight-residue sequence in the center was established. Repeat 3 was completely identified by sequencing a CNBr fragment. Whereas repeat 4 was not identified, repeat 5 was established except for the N-terminal residue by a chymotryptic and an elastase peptide.

C Terminus. Treatment of Galgp with carboxypeptidase Y released phenylalanine, asparagine, and leucine as major amino acids and valine, proline, and glycine as minor ones. From these data one might deduce a C-terminal peptide consisting of phenylalanine, asparagine, and leucine. However, the amino acid sequence does not contain a tripeptide or even a dipeptide of these residues. Hence, each of these amino acids probably represents the C terminus of a different molecular population. Since Phe-226 is the most N-terminal phenylalanine residue in the CD43 sequence, one species of Galgp likely extends up to residue 226, having the C-terminal sequence Val-Pro-Phe-COOH. This conclusion is supported by the finding that peptide PaPr-1 (Fig. 1) reaches from Pro-215 to Ser-222 as judged by its amino acid composition ($Asx_{1.0}Thr_{2.5}Ser_{2.1}Pro_{1.3}Ala_{0.9}Met_{0.5}$ in which Asx is aspartic acid or asparagine). Moreover, the composition corresponding to CD43 residues 1–226, $Asn_4Asp_4Gln_5Glu_{11}His_3Lys_5Arg_2Ala_{14}Gly_{18}Tyr_2Val_{13}Ile_6Leu_{12}Phe_1Met_7Trp_1Pro_{26}Ser_{45}Thr_{47}$ (no cysteine), agrees with the amino acid composition determined for purified Galgp (1, 3). Asparagine and leucine are probable C termini of further molecular populations of Galgp originating possibly from residues 228 or 232 and 193 or 209, respectively.

N-Glycan. Galgp possesses N-glycan (2) shown in this study to be attached to peptide PaPr-1, which contains Asn-220, the only site in the Galgp and CD43 amino acid sequence with the characteristic tripeptide, Asn-Xaa-(Ser or Thr), leading to N-glycosylation. Peptide PaPr-1 was shown earlier by NMR spectroscopic analysis to possess a fucosylated diantennary glycan (2). N-glycosylation has been demonstrated for human CD43 (6, 17), but the structure of the CD43 N-glycan has not been identified.

O-Glycans. The most unusual feature of Galgp is its high density of O-glycans. Previous studies estimate the number of Galgp O-glycans at 43 per molecule based on NMR spectroscopic analysis (2) and 57 per molecule based on compositional data and molecular weight (3). By sequencing both glycosylated and deglycosylated fragments in the present investigation, the glycosylation status was defined for 28 of the serine and threonine residues, of which 25 have O-glycans and 3 were shown to be nonglycosylated (Fig. 1). If this glycosylated to nonglycosylated ratio applies throughout the molecule, Galgp, which has 92 serine plus threonine residues, would contain ≈82 O-glycans. In the case of CD43, estimates based on compositional data (human and rat CD43; refs. 4–6) and peptide sequence (rat CD43; ref. 21) yield values of >75 O-glycans per molecule. NMR spectroscopic analysis (2) indicated that the predominant O-glycan of Galgp is a hexasaccharide with the following structure:

$$\left.\begin{array}{l}\text{NeuAc}(\alpha 2\text{–}3)\text{Gal}(\beta 1\text{–}3)\\ \text{NeuAc}(\alpha 2\text{–}3)\text{Gal}(\beta 1\text{–}4)\text{GlcNAc}(\beta 1\text{–}6)\end{array}\right\rangle \text{GalNAc}[\alpha 1\text{–}(\text{Ser or Thr})]$$

Naturally occurring CD43 species differ in the structure of their O-glycans, expressing predominantly tetrasaccharides in the case of circulating T lymphocyte and monocyte CD43 and the above hexasaccharide in the case of neutrophil, platelet, and activated T-lymphocyte CD43 (6,8,17,18).

Predicted Secondary and Tertiary Structures. The secondary structure of Galgp predicted from the primary amino acid sequence without consideration of O-glycans suggests predominantly random coil with some regions of β-sheet throughout the sequence. The structure of native Galgp measured by circular dichroic spectroscopy is characteristic of a predominantly random-coil somewhat-flexible conformation with ≈20% β-sheet (22).

As to its tertiary structure, the high density of O-glycans in Galgp is expected to exert a chain-stiffening and -lengthening effect because of steric interactions of the O-glycans, primarily the GalNAc residues, with the polypeptide (29, 32). Indeed, electron microscopy showed native Galgp to have extended filamentous shape with an average length of 43 nm and a particle width of about 4 nm (22). In the case of CD43, the molecule isolated from rat thymocytes was examined by electron microscopy, and the extracellular region was found to have an unfolded structure extending to a length of ≈45 nm (31) similar to the value of Galgp.

Mobility on SDS/PAGE and Immunoprecipitation. Purified Galgp was fractionated on SDS/PAGE and subjected to

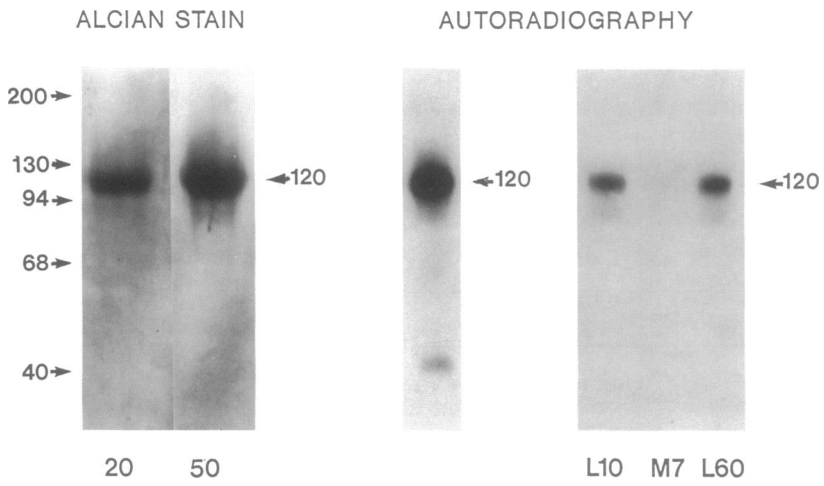

FIG. 3. Galgp analyzed by SDS/PAGE and immune precipitation. (*Left*) Galgp (20 μg, 50 μg) fractionated on SDS/PAGE (15, 23, 33) and stained with Alcian blue (34). The Galgp preparations (20 μg, 50 μg) did not stain with Coomassie blue (not shown). The Alcian blue-stained Galgp band has an apparent M_r of 120,000 relative to the Coomassie blue-stained molecular weight marker proteins (15, 23) indicated $\times 10^{-3}$ on the left. (*Middle*) Autoradiograph of Galgp (40 ng) labeled with ^{125}I by the iodogen method as described (5) and fractionated by SDS/PAGE. The major ^{125}I label is also found at an apparent M_r 120,000 (shown $\times 10^{-3}$). (*Right*) Autoradiograph showing immunoprecipitates prepared as described (15, 23) from 400 ng of ^{125}I-labeled Galgp by using the anti-CD43 monoclonal antibodies (mAbs) L10 and L60 and by the isotype-matched irrelevant mAb M7 (35).

staining with Coomassie blue and Alcian blue. Whereas the former dye detected no component, the latter dye revealed an apparent M_r 120,000 component (Fig. 3 *Left*). A dominant M_r 120,000 band was detected also when Galgp was labeled with ^{125}I and examined by autoradiography (Fig. 3 *Middle*). The apparent homogeneity suggests that the various Galgp species do not differ greatly in length. When examined by immunoprecipitation, ^{125}I-labeled Galgp was precipitated by two anti-CD43 mAbs with different specificities, L10 and L60, and was not precipitated by a subclass-matched irrelevant mAb (Fig. 3 *Right*). Immunoprecipitation of Galgp by anti-CD43 mAbs verifies the relatedness of these molecules.

Different molecular weight values have been reported previously for Galgp (1), which is not surprising since anomalous SDS/PAGE mobility is typical for highly glycosylated proteins. Nonetheless, the apparent value of M_r 120,000 in the present study can be contrasted with values of 115,000 for tetrasaccharide CD43 (CEM cell CD43) (8, 15) and 135,000 for hexasaccharide CD43 (neutrophil CD43) (8), since these were analyzed in parallel (not shown). These findings suggest that Galgp is more closely related to hexasaccharide CD43, consistent with the carbohydrate and compositional data.

Origin of Galgp. One explanation of the present findings—namely, the possibility that Galgp and CD43 are products of different genes—is largely excluded by earlier studies as follows. Southern analysis of human genomic DNA with a number of restriction enzymes and with cDNA probes encoding the common CD43/Galgp polypeptide region strongly indicates a single-copy gene or the product of a duplication event involving the entire gene and flanking region (19, 20). The possibility that Galgp and CD43 are generated by alternative splicing of exons of a common gene has also been largely excluded because the gene that encodes CD43 is devoid of detectable coding region introns (36). The extracellular, transmembrane, and cytoplasmic regions of CD43 are encoded by a single exon (36). A third possibility is that a single genomic sequence gives rise to two mRNA species, one encoding CD43 and one encoding Galgp, the latter generated by a process of RNA editing. RNA editing has been documented thus far only in the case of apolipoprotein B (37, 38).

Lastly, Galgp could be formed posttranslationally by proteolysis from transmembrane CD43. If it is generated by proteolysis, compositional data (1, 3) carbohydrate structure (2) and apparent molecular weight values suggest that Galgp is derived from "hexasaccharide CD43."

CONCLUSIONS

Amino acid sequencing of Galgp isolated from human plasma revealed a polypeptide identical to the extent examined (75%) with the N-terminal region of the blood-cell surface molecule CD43. The best characterized Galgp species consists of 226 amino acids, one N-glycan, and a high density of O-glycans (estimates range from 43 to 82 units per molecule). The expanded rod-like tertiary structure previously described for Galgp (22) is as expected for this mucin-like primary structure. Based on the common hexasaccharide structure of their O-glycans and on SDS/PAGE mobility, Galgp is most closely related to "hexasaccharide CD43," the species expressed on neutrophils, platelets, and T lymphocytes. The biological significance of Galgp, which is present in pooled plasma at levels greater than 10 μg/ml, is undefined, as is the origin of the molecule.

We are greatly indebted to Dr. Tomohiro Araki (Kyushu Tokai University, Kumamoto, Japan) for the C-terminal analysis, to Dr. Fumio Miake (The Chemo-Sero-Therapeutic Research Institute, Kumamoto, Japan) for the elucidation of the amino acid sequences of a large series of peptides and glycopeptides, to Scott Watkins (Department of Biology, Indiana University, Bloomington, IN) for computer evaluation of the amino acid sequence for secondary structure, to George Crombie for many amino acid analyses, to A. Steven Whitehead for helpful discussions, and to Abigail Allen for technical assistance. This study was supported by grants from the National Institutes of Health (GM-10374 to K.S., DK-43171 to M.A.H., AI29880 to E.R.-O., and HL-26335 to M.T.W.), the U.S. National Science Foundation (DMB-8510411 and DMB-8716413 to J.H.C.), the Deutsche Forschungsgemeinschaft (SFB-136 to R.B.), and the Swiss National Foundation (to J.S.).

1. Schmid, K., Mao, S. K. Y., Kimura, A., Hayashi, S. & Binette, J. P. (1980) *J. Biol. Chem.* **255**, 3221–3226.
2. Akiyama, K., Simons, R. E., Bernasconi, P., Schmid, K., van Halbeek, H., Vliegenthart, J. F. G., Haupt, H. & Schwick, H. G. (1984) *J. Biol. Chem.* **259**, 7151–7154.
3. Schwick, H. G. & Haupt, H. (1984) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, New York), pp. 166–202.
4. Brown, W. R. A., Barclay, A. N., Sunderland, C. A. & Williams, A. F. (1981) *Nature (London)* **289**, 456–460.

5. Remold-O'Donnell, E., Davis, A. E., III, Kenney, D., Bhaskar, K. R. & Rosen, F. S. (1986) *J. Biol. Chem.* **261,** 7526–7530.
6. Carlsson, S. R. & Fukuda, M. (1986) *J. Biol. Chem.* **261,** 12779–12786.
7. Axelsson, B., Hammarström, S., Finne, J. & Perlmann, P. (1985) *Eur. J. Immunol.* **15,** 427–433.
8. Remold-O'Donnell, E., Zimmerman, C., Kenney, D. & Rosen, F. S. (1987) *Blood* **70,** 104–109.
9. Silverman, L. B., Wong, R. C. K., Remold-O'Donnell, E., Vercelli, D., Sancho, J., Terhorst, C., Rosen, F., Geha, R. & Chatila, T. (1989) *J. Immunol.* **14,** 4194–4200.
10. Mentzer, S. J., Remold-O'Donnell, E., Crimmins, M. A. V., Bierer, B. E., Rosen, F. S. & Burakoff, S. J. (1987) *J. Exp. Med.* **165,** 1383–1392.
11. Nong, Y.-H., Remold-O'Donnell, E., LeBien, T. W. & Remold, H. G. (1989) *J. Exp. Med.* **170,** 259–267.
12. Vargas-Cortes, M., Axelsson, B., Larsson, A., Berzins, T. & Perlmann, P. (1988) *Scand. J. Immunol.* **27,** 661–671.
13. Remold-O'Donnell, E. & Rosen, F. S. (1990) *Immunodef. Rev.* **2,** 151–174.
14. Parkman, R., Remold-O'Donnell, E., Kenney, D. M., Perrine, S. & Rosen, F. S. (1981) *Lancet* **ii,** 1387–1389.
15. Remold-O'Donnell, E., Kenney, D. M., Parkman, R., Cairns, L., Savage, B. & Rosen, F. S. (1984) *J. Exp. Med.* **159,** 1705–1723.
16. Ardman, B., Sikorski, M. A., Settles, M. & Staunton, D. E. (1990) *J. Exp. Med.* **172,** 1151–1158.
17. Remold-O'Donnell, E., Kenney, D. & Rosen, F. S. (1987) *Biochemistry* **26,** 3908–3913.
18. Piller, F., Piller, V., Fox, R. I. & Fukuda, M. (1988) *J. Biol. Chem.* **263,** 15146–15150.
19. Shelley, C. S., Remold-O'Donnell, E., Davis, A. E., III, Bruns, G. A. P., Rosen, F. S., Carroll, M. C. & Whitehead, A. S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 2819–2823.
20. Pallant, A., Eskenazi, A., Mattei, M.-G., Fournier, R. E. K., Carlsson, S. R., Fukuda, M. & Frelinger, J. G. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1328–1332.
21. Killeen, N., Barclay, A. N., Willis, A. C. & Williams, A. F. (1987) *EMBO J.* **6,** 4029–4034.
22. Watzlawick, H., Walsh, M. T., Erhard, I., Slayter, H. S., Haupt, H., Schwick, H. G., Jourdian, G. W., Hase, S., Schmid, K. & Brossmer, R. (1991) *Biochem. J.* **277,** 753–758.
23. Remold-O'Donnell, E. & Rosen, F. S. (1990) *J. Immunol.* **145,** 3372–3378.
24. Stross, W. P., Warnke, R. A., Flavell, D. J., Flavell, S. U., Simmons, D., Gatter, K. C. & Mason, D. Y. (1989) *Clin. Pathol.* **42,** 953–961.
25. Gejyo, F., Chang, J.-L., Bürgi, W., Schmid, K., Offner, G. D. & Troxler, R. F. (1983) *J. Biol. Chem.* **258,** 4966–4971.
26. Araki, T., Gejyo, F., Takagaki, K., Haupt, H., Schwick, H. G., Bürgi, W., Marti, T., Schaller, J., Rickli, E., Brossmer, R., Atkinson, P. H., Putnam, F. W. & Schmid, K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 679–683.
27. Collins, J. H., Theibert, J. L., Francois, J.-M., Ashley, C. C. & Potter, J. D. (1991) *Biochemistry* **30,** 702–707.
28. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
29. Shogren, R., Gerken, T. A. & Jentoft, N. (1989) *Biochemistry* **28,** 5525–5536.
30. Carraway, K. L. & Hull, S. R. (1991) *Glycobiology* **1,** 131–138.
31. Cyster, J. G., Shotton, D. M. & Williams, A. F. (1991) *EMBO J.* **10,** 893–902.
32. Gerken, T. A., Butenhof, K. J. & Shogren, R. (1989) *Biochemistry* **28,** 5536–5543.
33. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
34. Wardi, A. H. & Michos, G. A. (1972) *Anal. Biochem.* **49,** 607–609.
35. Remold-O'Donnell, E. & Savage, B. (1988) *Biochemistry* **27,** 37–41.
36. Shelley, C. S., Remold-O'Donnell, E., Rosen, F. S. & Whitehead, A. S. (1990) *Biochem. J.* **270,** 569–576.
37. Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J. & Scott, J. (1987) *Cell* **50,** 831–840.
38. Higuchi, K., Hospattankar, A. V., Law, S. W., Meglin, N., Cortright, J. & Brewer, H. B., Jr. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 1772–1776.