Human Keratinocytes Express a New CD44 Core Protein (CD44E) as a Heparan–Sulfate Intrinsic Membrane Proteoglycan with Additional Exons

Tod A. Brown,* Todd Bouchard,* Tom St. John,* Elizabeth Wayner,‡ and William G. Carter*‡§

*Fred Hutchinson Cancer Research Center and §Department of Pathobiology, University of Washington, Seattle, Washington 98104; ‡Cytel Corporation, La Jolla, California 92037

Abstract. We previously identified a 90-kD (GP90), collagen-binding, membrane glycoprotein, termed extracellular matrix receptor III (ECMR III), that is homologous to the lymphocyte homing receptor and CD44 antigen (Gallatin, W. M., E. A. Wayner, P. A. Hoffman, T. St. John, E. C. Butcher, and W. G. Carter. 1989. *Proc. Natl. Acad. Sci. USA.* 86:4654–4658). CD44 is abundantly expressed in many epithelial tissues, and is localized predominantly to filopodia in cultured keratinocytes. Here we establish CD44 as a polymorphic family of related membrane proteoglycans and glycoproteins possessing extensive diversity in both glycosylation and core protein sequence. Human neonatal foreskin keratinocytes (HFKs) and QG56 lung squamous carcinoma cells express an alternatively spliced form of the CD44 core protein (termed CD44E) that contains an additional 132 amino acids in the carbohydrate attachment region of the extracellular domain. HFKs, HT1080 fibrosarcoma and QG56 cells, as well as many other human cells, contain varying ratios of GP90 and structurally related, higher molecular mass forms of CD44 that express the following characteristics: (*a*) each form

reacted with anti-CD44 (mAbs) P1G12, P3H9, and P3H5. Each of these mAbs recognized a distinct, nonoverlapping epitope present on each CD44 form. (*b*) Differences in mass were due primarily to variation in carbohydrate moieties, including sulfated aspargine-linked glycopeptides (GP), chondroitin sulfate (CS), and heparan sulfate (HS) glycosaminoglycans, as well as O-linked mucin and polylactosamine structure(s). The major polymorphic forms were designated HT1080 GP90 and CS180, QG56 GP230, and HFK HS/CS250, based on dominant carbohydrate moieties and relative mass. (*c*) The polymorphic forms use CD44 and CD44E core proteins, each containing a unique set of potential attachment sites for O- and N-glycosides and glycosaminoglycans. (*d*) Immunofluorescence microscopy, differential extraction with Triton-X-114 detergent, and incorporation into liposomes indicated that all the forms were membrane bound glycoconjugates. These results define CD44 as a structurally diverse, but immunologically related, set of intrinsic membrane macromolecules, and suggests that these structurally varied forms might be expected to manifest multiple functions.

THE polymorphic family of integral membrane glycoproteins CD44 (see Haynes et al., 1989 for review) is found on a wide variety of cells. The exact function(s) of CD44 has yet to be conclusively defined, although recent reports have implicated CD44 in lymphocyte homing (Jalkanen et al., 1987), lymphohemopoiesis (Miyake et al., 1990*a*) and T cell activation (Shimizu et al., 1989; Huet et al., 1989). The widespread distribution of CD44 in nonhemopoietic cell types (Carter and Wayner, 1988; Flanagan et al., 1989), however, argues for a more general function than those delegated to it in the hemopoietic lineage. Previously reported biochemical and cell biological properties are consistent with a role as an adhesive cell-surface glycoprotein. CD44 may function either by linking extracellular matrix proteins such as collagen (Carter and Wayner, 1988) or

hyaluronic acid (Miyake et al., 1990*b*; Aruffo et al., 1990) with the cytoskeleton (Carter and Wayner, 1988; Kalomiris and Bourguignon, 1988), or may be involved with homotypic (St. John et al., 1990; Belitsos et al., 1990) or heterotypic (Shimizu et al., 1989; King et al., 1990) intercellular adhesion.

Most studies have concentrated on the 80–90-kD form of CD44 in hemopoietic cells. Before the CD44 classification, it was referred to by several designations, including extracellular matrix receptor III (ECMR III; Wayner and Carter, 1987; Carter and Wayner, 1988), phagocytic glycoprotein 1 (Pgp-1; Isacke et al., 1986), the hyaluronate receptor (Alho and Underhill, 1989; Underhill, 1989), p80 (Haynes et al., 1983), p85 (LeTarte et al., 1985), or antigens of the mAbs F-10-44-2 (Dalchau et al., 1980), Hermes-1 (Jalkanen et al.,

1987), or Hutch-1 (Gallatin et al., 1989). Recent studies using a host of monoclonal antibodies have verified immunologically that these molecules are very similar, if not identical, entities (Gallatin et al., 1989; Picker et al., 1989; Omary et al., 1988; Pals et al., 1989; Lesley et al., 1990). The predominant mRNA transcripts of CD44 in baboon (Idzerda et al., 1989) and human (Stamenkovic et al., 1989; Goldstein et al., 1989) lymphoid cells codes for a 37-kD core protein with 75, 5, and 20% of the polypeptide designated as extracellular, transmembrane, and cytoplasmic domains, respectively. The NH_2-terminal portion of the extracellular domain has an ∿30% homology to mouse link protein, a hyaluronic acid and collagen binding protein which stabilizes large aggregates of proteoglycan core proteins in cartilage (Neame et al., 1987). This protein core of CD44 is modified by extensive N- and O-glycosylation to an 80–90-kD form or by addition of chondroitin sulfate to a 180–200-kD form (Carter and Wayner, 1988; Gallatin et al., 1989; Jalkanen, 1988; Stamenkovic et al., 1989; Omary et al., 1988).

In this report we have investigated the structural aspects of the higher molecular mass forms, particularly in nonhematopoietic cells. We characterized four of these forms from three different cell populations, with particular attention to human neonatal foreskin keratinocytes (HFKs),[1] a cell population that expresses particularly high quantities of CD44 in tissue (Carter et al., 1990a). The major form of CD44 expressed by HFKs, termed HS/CS250, is identified as a membrane heparan sulfate/chondroitin sulfate proteoglycan and was localized to filopodia and intercellular contacts.

Materials and Methods

Materials

PMSF, N-ethylmaleimide, 2-mercaptoethanol, BSA, Triton X-100 and Triton X-114 detergents, n-octyl-β-D-glucopyranoside detergent, ethylenediaminetetraacetic acid, protein A-agarose, egg yolk phosphatidylcholine, were purchased from Sigma Chemical Co. (St. Louis, MO). Random hexamer and oligo dT_{12} were obtained from Pharmacia Fine Chemicals (Piscataway, NJ). Pronase and lactoperoxidase were from Calbiochem-Behring Corp. (La Jolla, CA). Fluorescein-conjugated (goat) anti-mouse IgG and IgM (H and L chains) and rhodamine-conjugated (goat) anti-rabbit IgG were obtained from Tago, Inc. (Burlingame, CA). Rabbit anti-mouse IgG (H+L) antiserum was obtained from Cappel Laboratories (Cooper Biomedicals, Malvern, PA). [^{35}S]H_2SO_4 (carrier-free) and [^{35}S]methionine (Trans ^{35}S-label, 1,010 Ci/mmol) were from ICN Radiochemicals (Cleveland, OH). Enlightning, [α-^{32}P]deoxycytidine triphosphate (3,000 Ci/mmol) and ^{3}H-phosphatidylcholine were purchased from New England Nuclear (Wilmington, DE). ^{125}I and [α-^{35}S]deoxyadenosine thiotriphosphate (>1,000 Ci/mmol) was from Amersham Chemical Co. (Arlington Heights, IL). Wheat germ agglutinin (WGA) was obtained from Vector Laboratories (Burlingame, CA) and was coupled to CNBr-activated Sepharose (Parikh et al., 1974). Glycosidases and DNA-modifying enzymes, unless noted elsewhere, were obtained from Boehringer Mannheim Biochemicals (Indianapolis, IN).

Cells, Cell Culture, Tissue

Normal newborn HFKs were prepared as described by Boyce and Ham (1985). Cultures were maintained in serum-free keratinocyte growth medium (KGM; Clonetics, San Diego, CA). Normal embryonic human lung diploid fibroblasts, WI-38 cells, BT-20 human breast carcinoma, and a hu-

1. *Abbreviations used in this paper*: HFF, human dermal fibroblasts; HFK, human foreskin fibroblasts.

man fibrosarcoma, HT1080, were obtained from the American Type Culture Collection (Rockville, MD). QG-56 (squamous cell carcinoma of the lung) was obtained from Dr. Sen-itiroh Hakomori (The Biomembrane Institute, Seattle, WA). These cell lines were grown as adherent monolayer cultures in RPMI 1640 supplemented with 10% FBS.

Human lymphokine-activated killer (LAK) cells were grown from peripheral blood lymphocytes (PBL) as described (Grimm et al., 1982) in Click's medium supplemented with 10% FBS and 500–1,000 U/ml IL2. All other lymphoid tumor cell lines: Molt 4 (T cell leukemia), HL-60 and U937 (promonocyte cell lines), and Daudi (Burkitt's lymphoma) were obtained from Dr. David Urdal (Immunex Corp., Seattle, WA). Primary spleen cells, the EBV transformed human B lymphocyte cell line ST-1, and PHA stimulated T lymphocytes were prepared as previously described (Wayner and Carter, 1987; Wayner et al., 1988). CLL is a chronic T lymphocyte leukemia derived from a patient in blast crisis. Manka cells are a human Burkitts lymphoma.

Preparation of Monoclonal and Polyclonal Antibodies

Anti-CD44 mAbs (P3H9, P3H5, and P1G12) were prepared as previously described (Wayner and Carter, 1987). mAbs P1G12, P3H9, and P3H5 all reacted with nonoverlapping epitopes on CD44 as determined by competitive binding studies (Gallatin et al., 1989).

Rabbit polyclonal antibodies against human fibronectin were prepared as previously described (Carter, 1982). Rabbit polyclonal antibodies were prepared against mouse laminin isolated from mouse EHS sarcoma. Rabbit polyclonal antiserum against involucrin, a component of the keratinocyte envelope (Rice and Green, 1979), was a gift from Dr. Robert Rice (University of California, Davis, CA). Rabbit polyclonal antibodies (Ab33; Roman et al., 1989) against integrin α5 were from Dr. John McDonald (Washington University, St. Louis, MO). Rabbit polyclonal antiserum against type I collagen was from Chemicon International, Inc. (El Segundo, CA).

Metabolic Labeling, Pulse–Chase Experiments, and PAGE

Metabolic labeling with $^{35}SO_4^{-2}$ or with [^{35}S]methionine or pulse-chase studies using [^{35}S]methionine were performed essentially as described previously (Carter and Wayner, 1988).

Polyacrylamide slab gels containing SDS (SDS-PAGE gels) were prepared following the basic stacking gel system of Laemmli (1970). Prestained protein standards for relative molecular mass estimation were obtained from either Bethesda Research Laboratories (Bethesda, MD) or Sigma Chemical Co. (St. Louis, MO). Protein concentration was determined by the fluorescamine method using BSA as standard (Udenfriend et al., 1972).

Nitrous Acid Cleavage of Heparan Sulfate

Triton X-100 extracts of $^{35}SO_4^{-2}$-labeled HFKs or QG56 cells were immune precipitated with P3H9 to isolate radiolabeled CD44. The selective removal of heparan sulfate by the method of Shively and Conrad (1976) was performed as described by David and Van den Berghe (1985), except incubations were performed for 0, 15, or 60 min. Cleavage products were analyzed by SDS-PAGE. The selectivity of this method was confirmed by the insensitivity of the QG56 GP230 form (which lacks significant heparan sulfate addition) to cleavage relative to the HFK HS/CS250 form.

Immunofluorescence, Interference Reflection, and Antibody-Exclusion Microscopy

Cells were adhered to coverslips as described (Carter et al., 1990a) and were incubated with combinations of mouse or rat mAbs and rabbit polyclonal primary antibodies and processed for visualization by two-color fluorescence, interference reflection or antibody exclusion microscopy as previously described (Carter et al., 1990a). Interference reflection microscopy (IRM) was performed basically as described (Izzard and Lochner, 1976) and was used to identify focal adhesions (FAs) in the same field as the two color fluorescence. FAs were also localized by the antibody exclusion technique (Neyfakh et al., 1983; Carter et al., 1990a).

Fluorescence Microscopy of Tissue Sections

The distribution of receptors and ligands in tissue were determined by immunofluorescence microscopy of 6-μm cryostat sections as previously

described (Wayner et al., 1988). All sections were fixed in 4% paraformaldehyde in PBS before incubation in primary antibodies and peroxidase-conjugated secondary antibodies. The samples were examined with a Zeiss fluorescence microscope and photographed with Kodak T-Max film at ASA 800. In control experiments, no fluorescence of rhodamine was detected using the fluorescein filters or vice versa.

PCR Amplification

The isolation of poly-A$^+$ RNA from primary HFKs, dermal fibroblasts (HFFs), and the human cell lines HT1080 or QG56 was as described by Badley et al. (1988). Two micrograms of poly-A$^+$ RNA was transcribed into cDNA by reverse transcriptase using either random hexamer oligomers or oligo (dT)$_{12}$ as primers and cDNAs were purified by gel exclusion chromatography (BioGel A5m; Bio-Rad Laboratories, Richmond, CA). Purified cDNA was amplified by polymerase chain reaction (PCR) in a DNA thermal cycler (Perkin-Elmer Cetus Corp., Norwalk, CT) using *Thermus aquaticus* (*Taq*) DNA polymerase and oligonucleotide pairs as noted in the text (see Fig. 8 *A* for an oligonucleotide map). Other buffer components were as described by the enzyme supplier (Perkin-Elmer Cetus Corp.). Reaction conditions were as follows: 4 min at 94°C followed by 35 cycles of 1 min at 94°C, 2 min at 50°C and 4 min at 72°C. The reaction products obtained by PCR were separated by agarose gel electrophoresis; individual bands were excised and the cDNA isolated by the freeze/squeeze method (Tautz and Renz, 1983).

DNA Sequencing and Analysis

Double-stranded cDNA for sequencing was obtained by a second round of PCR amplification using the purified cDNA from the first round as template. DNA was purified by gel electrophoresis, as described above, and used directly for sequencing by the dideoxy chain termination method (Sanger et al., 1977), using a modified T7 DNA polymerase (Sequenase; U. S. Biochemical Corp., Cleveland, OH) and the modification of Winship (1989). Primers were annealed to the DNA by snap cooling from 100°C

into dry ice, and included D1, B7-6, HE1, and SJ1 oligonucleotides (see Fig. 8 *A*). Both strands from at least three independent, primary PCR reactions were amplified and sequenced to check and correct for possible spontaneous mutations caused by *Taq* DNA polymerase (Ennis et al., 1990). DNA sequence analysis was performed with Genepro software (Riverside Scientific, Seattle, WA).

Northern Blot Analysis

Northern blot analysis of formaldehyde-treated, size fractionated poly-A$^+$ RNA was, with slight modification, as described by Selden (1989). Hybridization probes were either derived from cDNA clones or synthesized by PCR, tailored for use (if necessary) by restriction digestion and subsequently purified by elution of the appropriate band following agarose gel electrophoresis. Probes were labeled with ^{32}P by extension of random hexamer primers using Klenow fragment of DNA polymerase I in the presence of [alpha-^{32}P]dCTP (Feinberg and Vogelstein, 1983). Molecular sizes were determined relative to ethidium bromide-stained molecular mass markers (Bethesda Research Laboratories).

Results

Expression of CD44 in Epithelial Tissue and Cultured Cells

Localization of CD44 in cryostat sections of human tissue, particularly epithelium, identified extensive expression in many nonhematopoietic cells. In neonatal foreskin (Fig. 1), CD44 was preferentially expressed around the entire periphery of epidermal cells in the basal and spinous layers. CD44 was differentially expressed and decreased in the stratified layers subsequent to expression of involucrin, a
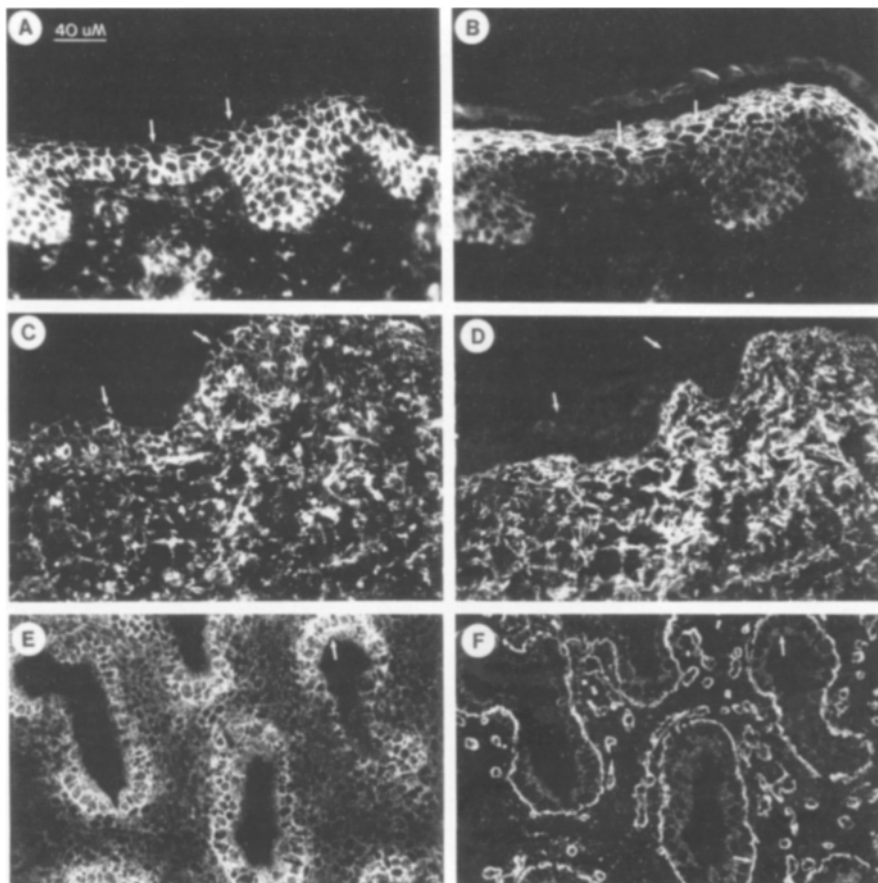

Figure 1. Localization of CD44 on basal, lateral, and apical surfaces of epithelial cells in cryostat sections of human neonatal foreskin, fetal skin, and fetal lung. (*A* and *B*, same field, foreskin) Anti-CD44 (*A*, mAb P3H9) stains entire epidermal cell periphery in basal, and spinous layers but is absent from granular layers that stain with anti-involucrin (*B*). Arrows identify the same points in the two fields and distinguish the staining of CD44 and involucrin. (*C* and *D*, same field, fetal skin) Anti-CD44 (*C*) in epidermis and dermis while anti-fibronectin (*D*) localizes only in dermis. (*E* and *F*, same field, fetal lung) Anti-CD44 (*D*) strong staining in developing bronchiol epithelium that is surrounded by laminin (*F*) in the basement membrane. Bar: (*A*–*F*) 40 μm.
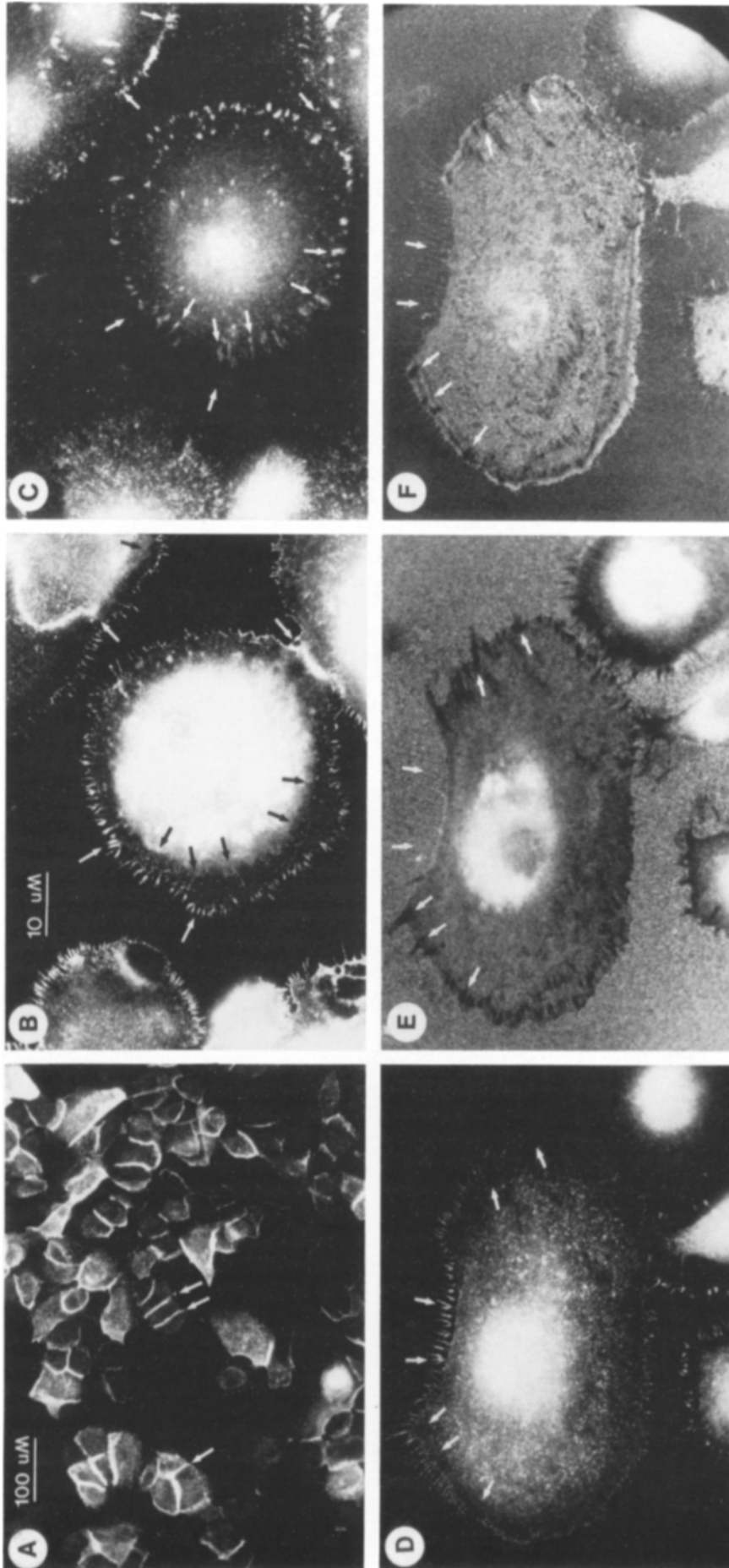
Figure 2. Localization of CD44 (HS/CS250) in filopodia and cell–cell contacts of cultured HFKs. HFKs were adhered to fibronectin- or type I collagen-coated coverslips and then the indicated receptors were localized by immunofluorescence microscopy. Where indicated, focal adhesions were identified by interference reflection microscopy or antibody exclusion techniques (see Materials and Methods). (*A*) CD44 localized in Ca²⁺-induced HFKs with mAb P3H9. Arrows indicate accumulations of CD44 at cell–cell contacts. (*B*) HFKs adhered to fibronectin. CD44 accumulates in filopodia on apical surface and at points of cell contact with the substrate (*white arrows*). (*C*, same field as *B*) Integrin α5β1 (fibronectin receptor) localized in focal adhesions. Arrows in *B* and *C* designate the same points and distinguish CD44 in filopodia and α5β1 in focal adhesions. (*D*) CD44 in filopodia at periphery of HFKs adhered to type I collagen. (*E*, same field as *D*) Focal adhesions (dark hole designated by *arrows*) localized by staining with anti-type I collagen (antibody exclusion technique). (*F*, same field as *D* and *E*) Focal adhesions identified by interference reflection microscopy. CD44 locates near but not in the focal adhesions. Bars: (*A*) 100 μm; (*B–F*) 10 μm.
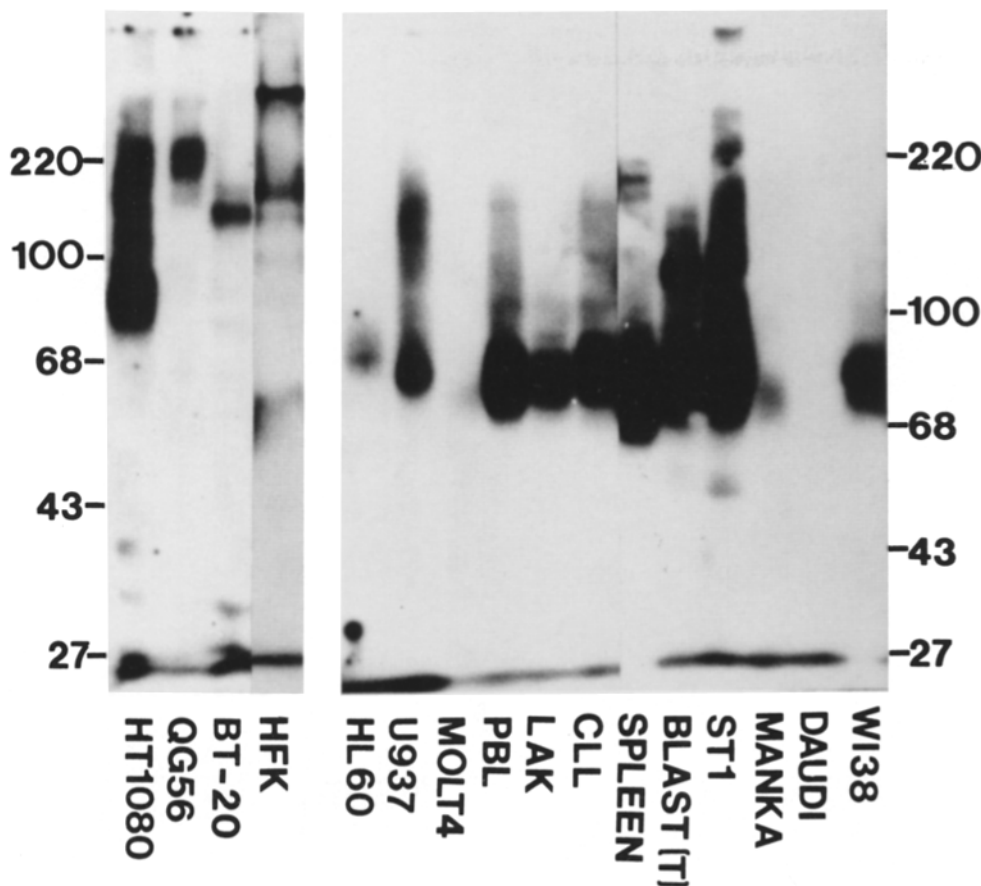
Figure 3. Polymorphism in CD44 expressed by different cell populations. Triton X-100 detergent extracts of the indicated cell populations (HT1080, QG56, etc., 75 µg protein per cell extract), were fractionated by SDS-PAGE and immunoblotted with monoclonal antibody P1G12 against CD44.

component of the keratinocyte envelope (Rice and Green, 1979; compare Fig. 1, *A* and *B*). In fetal skin (Fig. 1 *C*), CD44 was expressed in both the epidermis and dermis, whereas fibronectin was expressed in only the dermis (Fig. 1 *D*). In human fetal lung (Fig. 1 *E*), CD44 was also expressed around the entire periphery of cells in developing bronchial epithelium surrounded by the laminin containing basement membrane (Fig. 1 *F*).

CD44 antigens, detected by immunofluorescent microscopy in cultured keratinocytes, were localized to the cell surface in HFKs (Fig. 2). Although CD44 was distributed over the entire cell surface, it was clearly increased in concentration in filopodia present on the apical surface of cells and at the periphery of cells where they contact the substrate (Fig. 2 *B*). The concentration in filopodia, however, was distinct from stable cell attachment sites such as focal adhesions (Burridge et al., 1988; Carter et al., 1990a). CD44 was not detected in focal adhesions of cells attached to collagen, laminin or fibronectin. In contrast, integrins $\alpha5\beta1$ (Fig. 2 *C*) and $\alpha2\beta1$, $\alpha3\beta1$ (Carter et al., 1990a,b) were readily detected in elevated concentrations in focal adhesions of many cells. Negligible CD44 was found in the extracellular matrix.

In cultures of HFK cells that had been induced to aggregate by growth in the presence of Ca^{2+}, CD44 concentrated in sites of cell–cell contact (Fig. 2 *A*), suggesting that it may be involved in cell–cell adhesion. This function would be consistent with the known relationship between CD44 and homotypic cell aggregation.

Identification of Polymorphic Forms of CD44

We have previously identified three mAbs, P1G12, P3H9,

and P3H5 that react with distinct epitopes on the GP90 form of CD44 antigen from HT1080, human fibrosarcoma cells (Carter and Wayner, 1988). We have used these mAbs in immunoblotting experiments to establish that different cell populations express one or more distinct molecular mass forms of CD44, ranging from 80 to >250 kD (Fig. 3). All three mAbs identified the same forms of CD44 in each cell population, indicating that all the forms expressed the same nonoverlapping epitopes.

All the forms detected by immunoblotting or in subsequent immunoprecipitation experiments migrated as broad diffuse bands on SDS-PAGE suggesting a high degree of posttranslational modification, such as phosphorylation and/or glycosylation (Carter and Wayner, 1988): (*a*) HT1080 fibrosarcoma, U937 premonocyte and ST-1 Epstein–Barr virus-transformed B lymphocytes all expressed at least two forms of CD44 migrating with relative molecular masses of ∼90 and 180 kD. In general the 180-kD form was always present in minor quantities relative to the 90-kD form. (*b*) QG56 lung carcinoma (230 kD), neonatal HFKs (180 and 250 kD), and BT-20 breast carcinoma (120 kD), expressed the indicated higher molecular mass forms as the predominant species, with relatively lower quantities of the 90-kD forms detected. In contrast some cells such as Molt 4, T cell lymphoma, and Daudi, a Burkitts lymphoma expressed no detectable CD44 under these conditions.

Characterization of Sulfated and Nonsulfated Carbohydrates on CD44

We investigated carbohydrate diversity as a possible source of CD44 polymorphism. First, we examined the CD44 anti-
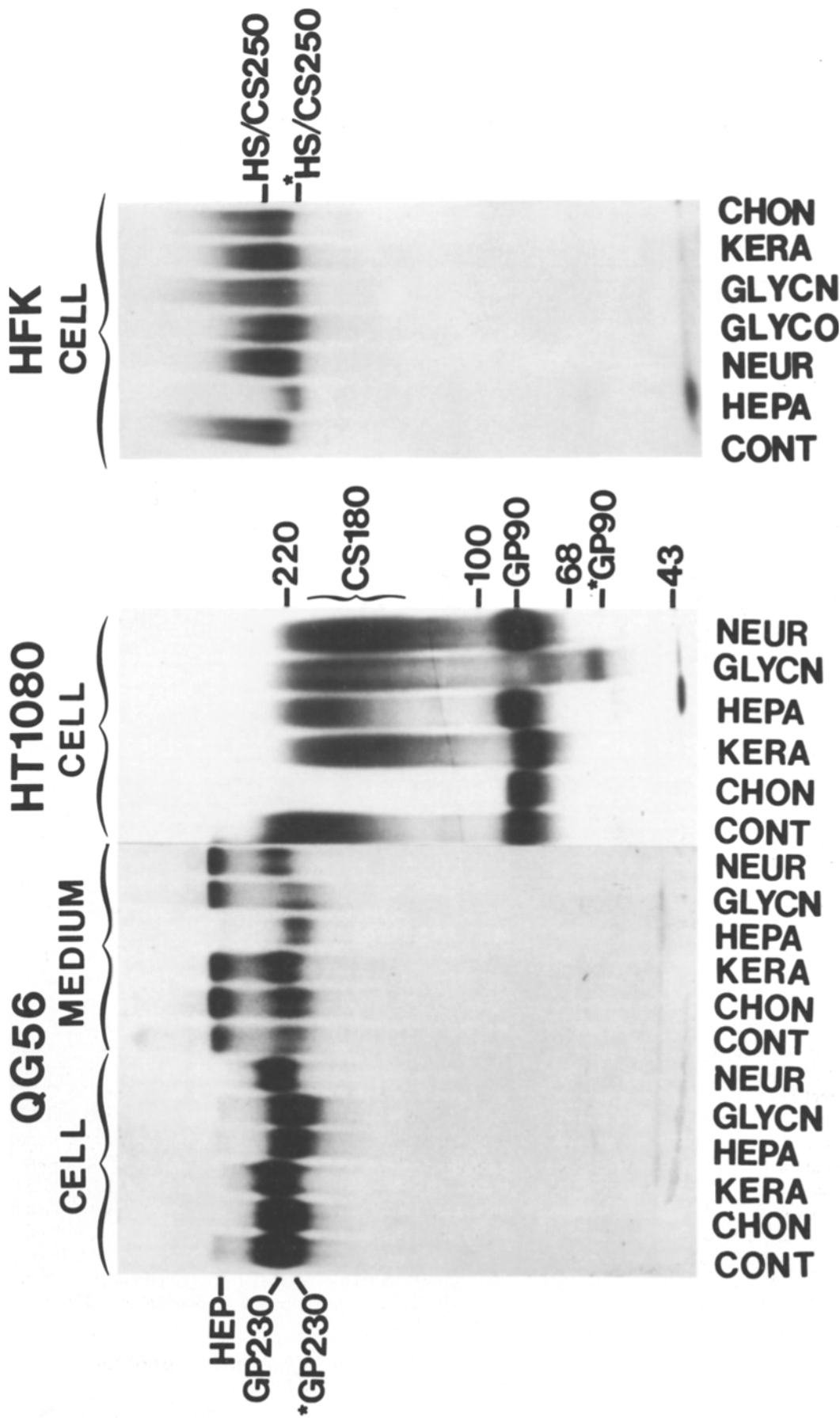
Figure 4. Metabolic labeling of HT1080, HFK, and QG56 cells with $^{35}SO_4^{-2}$, immunoprecipitation and digestion of polymorphic CD44 antigens with carbohydrate degrading enzymes. QG56, HFK, and HT1080 cells were metabolically labeled with $^{35}SO_4^{-2}$, extracted with Triton X-100 detergent and CD44 was immunoprecipitated from the extracts with mAb PlGl2 yielding GP90 and CS180 from HT1080 cells, HS/CS250 from HFKs, and GP230 from QG56 cells and conditioned culture medium of QG56 cells. The purified polymorphic forms of CD44 were digested with protease-free carbohydrate degrading enzymes: chondroitin ABC lyase (*CHON*, from *Proteus vulgaris*, degrades chondroitin sulfate or hyaluronic acid); keratan sulfate, 1,4-β-D-galactanohydrolase (*KERA*, from Pseudomonas); heparitinase (*HEPA*, from *Flavobacterium heparinum*); glycopeptide-*N*-glycosidase (*GLYCN*, from *Flavobacterium meningosepticum*); neuraminidase (*NEUR*, from *Vibrio cholera*); and neuraminidase, followed by glycopeptide-*o*-glycosidase (*GLYCO*, from *Diplococcus pneumoniae*). The digested antigens were fractionated by SDS-PAGE (8%) and detected by fluorography. Migration of molecular mass standards are indicated. Major degradation products exhibiting greater mobility, but retaining $^{35}SO_4^{-2}$, are identified with asterisks.
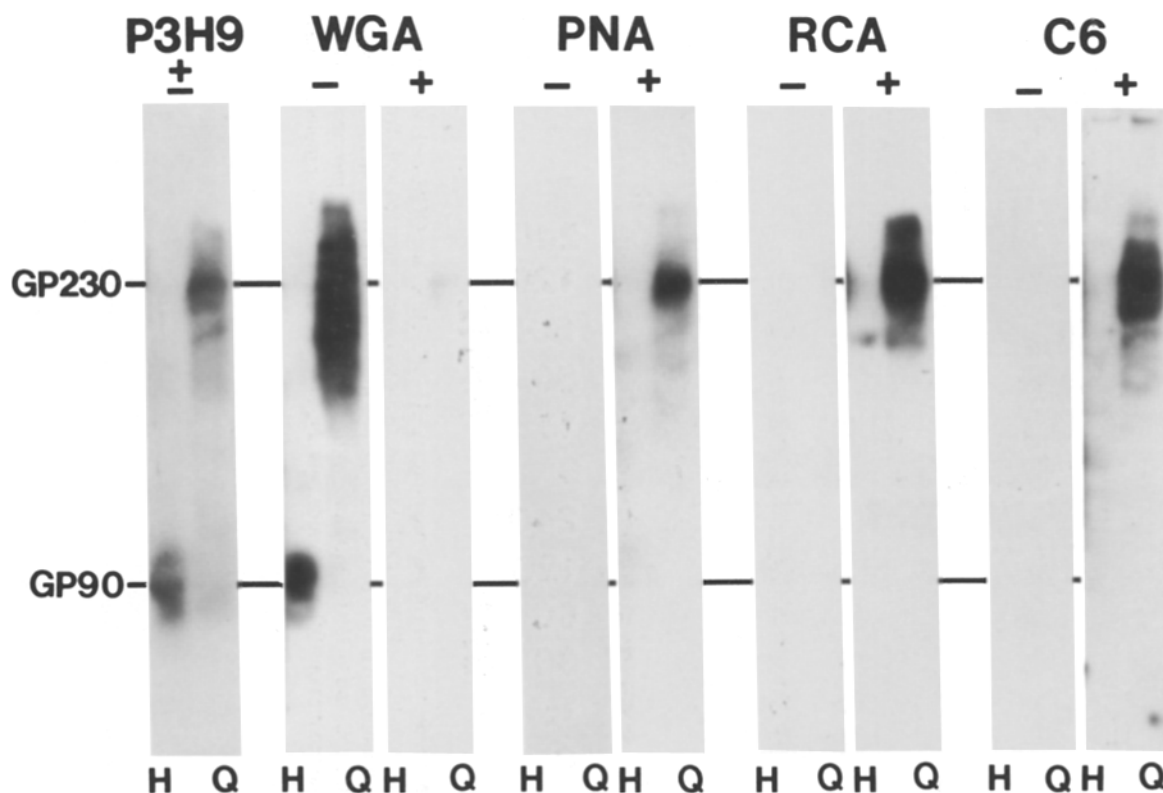
Figure 5. Effects of neuraminidase digestion on lectin binding to HT1080 GP90 and QG56 GP230. Purified GP90 from HT1080 cells and GP230 from QG56 cells were fractionated by SDS-PAGE (8%) then electrophoretically transferred to nitrocellulose. The nitrocellulose-bound glycoproteins were incubated with (+) or without (−) neuraminidase, then incubated with ^{125}I-labeled lectins or antibodies followed by ^{125}I-labeled protein A. Antibodies and lectins bound to the glycoproteins were detected by autoradiography.

gens for the presence of high molecular mass sulfated carbohydrates. The presence of extensive sulfation on the multiple forms of CD44 was determined by metabolic labeling of HT1080, QG56 and HFK cells with $^{35}SO_4^{-2}$ followed by immunoprecipitation of the labeled cell extracts with mAb P1G12 (Fig. 4). To facilitate identification of the forms of CD44, we referred to each form on the basis of the relative molecular mass and the dominant carbohydrate moieties detected by enzyme digestion. The results are summarized as follows: (*a*) digestion of the 90-kD form from HT1080 cells (Fig. 4, *HT1080*) with glycopeptide-*N*-glycosidase removed ~50% of the sulfate label associated with the 90-kD form resulting in a higher mobility form labeled *GP90. None of the other enzymes had a significant effect on the degree of sulfation and/or migration. Therefore, the only detectable form of carbohydrate sulfation associated with the 90-kD form is sulfated asparagine-linked glycopeptides (GP). This form shall be referred to as HT1080 GP90. Unsulfated, *O*-glycosidically linked mucin carbohydrates were also detected on GP90 by digestion with glycopeptide-*O*-glycosidase, after neuraminidase treatment, and a corresponding increase in migration of the GP90 resulted (not shown). The presence of enzyme-resistant sulfate on both HT1080 GP90 and QG56 GP230 suggests possible peptide sulfation. (*b*) Sulfate associated with HT1080 (180-kD form) was completely removed by digestion with chondroitinase ABC but was unaffected by the other enzymes. Thus, all the sulfate is in the form of chondroitin sulfate (CS) and this form shall be referred to as HT1080 CS180. (*c*) None of the enzymes, either individually (Fig. 4, *QG56*) or in combination (results not shown) removed >60% of the sulfate from the 230-kD form in QG56 cells. However, both heparitinase and glycopeptide-*N*-glycosidase released ~20–30% of the label from the core protein with a corresponding increase in the band migration, labeled *GP230. This suggests that the 230-kD form contains a minor percentage of the covalently bound sulfate in the form of heparan sulfate and sulfated asparagine-linked carbohydrates and shall be referred to as QG56 GP230. Immune precipitation of CD44 from conditioned culture medium of QG56 (Fig. 4, *QG56, Medium*) and HFKs (not shown), but not HT1080 cells, detected released-forms of the CD44 antigens. The released forms comigrated with the cellular forms of the antigens from each cell type. Enzyme digestion of the released forms gave similar results as the cellular forms. The released form of GP230 contained a high molecular weight component, designated "HEP," that coprecipitated with GP230 and was completely degraded by the heparitinase. Various dissociation experiments failed to remove HEP from the immunoprecipitates, suggesting that it is a higher molecular mass form of CD44 with extensive heparan sulfate additions. (*d*) Most of the sulfate associated with the 250-kD forms of CD44 from HFK cells was removed by digestion with heparitinase (Fig. 4, *HFK, Cell*) or by a 15-min incubation with nitrous acid (not shown, Shively and Conrad, 1976). The minor amount of label associated with the heparitinase-resistant core was removed by diges-
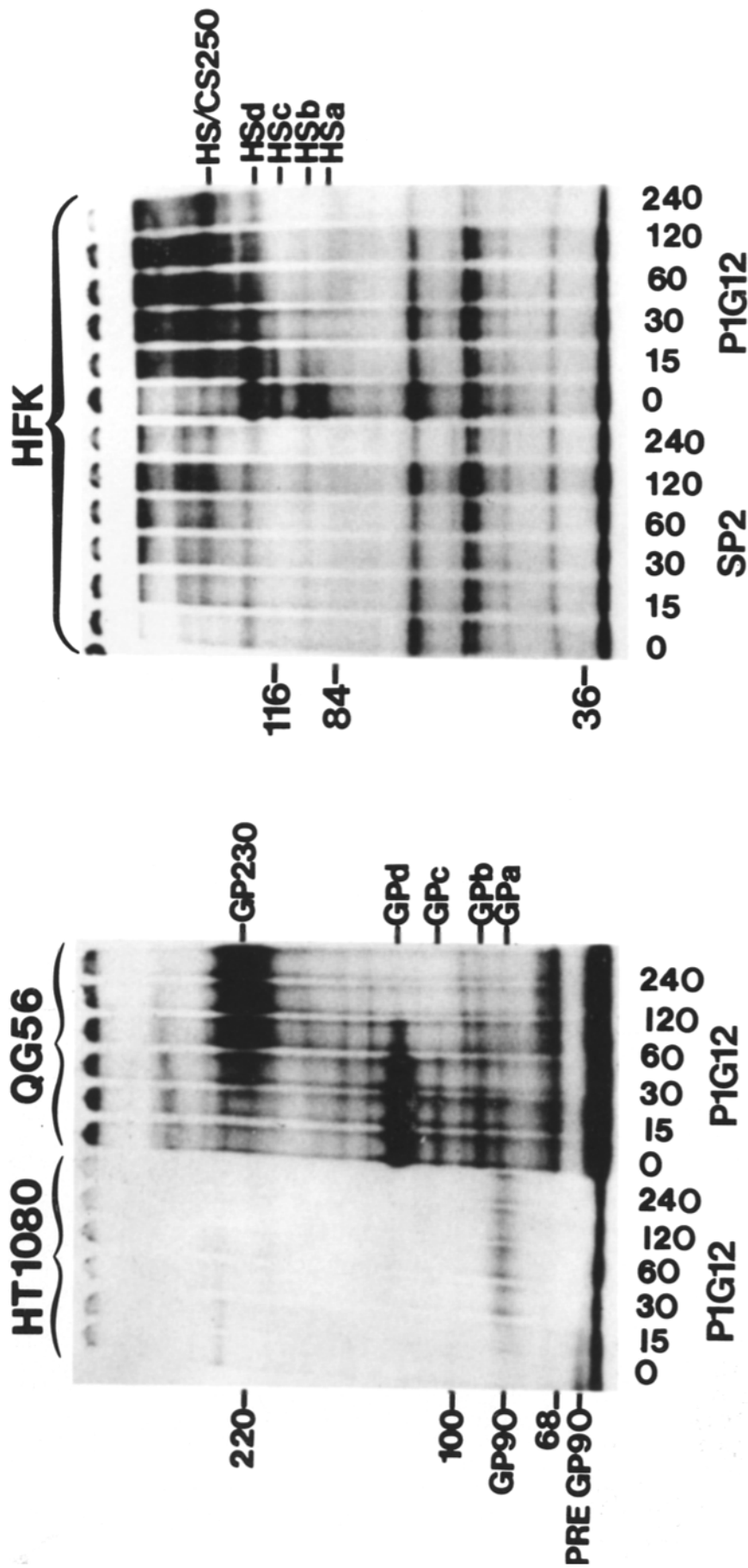
Figure 6. Pulse–chase metabolic labeling of HT1080 GP90 and QG56 GP230, and HFK HS/CS250 with [³⁵S]Met: identification of the core proteins. HT1080, HFK, and QG56 cells were pulse-labeled with [³⁵S]met for 30 min then chased in the presence of unlabeled met. At the end of the indicated time periods (0, 15, 30 min, etc.), samples of the cells were extracted with Triton X-100. The extracts were immune-precipitated with either control culture supernatant (*SP*) or culture supernatant containing P1G12 mAb. The immune-precipitates were fractionated on SDS-PAGE gels (7%, HT1080 and QG56; 10%, HFK) followed by fluorography. Migration of HT1080 GP90, HFK HS/CS 250, and QG56 GP230 along with the major precursor core proteins, preGP90 (65 kD), *HSa, b, c,* and *d,* and *GP a, b, c,* and *d* (90–135 kD), respectively, are indicated. Analysis of duplicate samples on 10% acrylamide gels also failed to detect any lower molecular mass precursors for either GP90 or GP230.
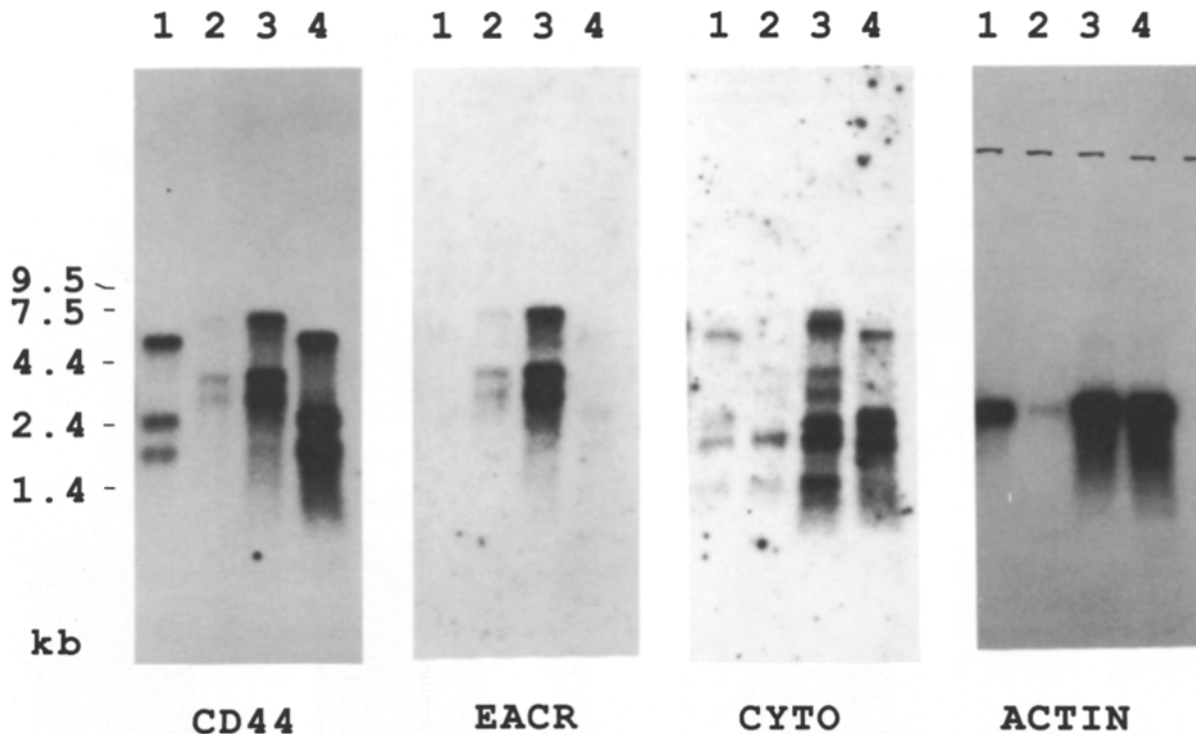
214

Figure 7. Northern blot analysis with different CD44 probes using poly-A⁺ RNA from four cell populations. Poly-A⁺ RNA from HT1080, QG56, HFK, and HFF cells (lanes *1-4*, respectively) was size fractionated by agarose gel electrophoresis, transferred to nitrocellulose and probed sequentially with a full-length lymphoid CD44 cDNA from baboon (93% homologous to the human lymphoid sequence, *CD44*), a hamster actin probe (*ACTIN*), an EACR-specific cDNA (*EACR*), and a CD44 cytoplasmic domain cDNA (*CYTO*). The EACR-specific cDNA was a 372-bp fragment corresponding to nucleotides 803-1172 of the reported sequence. It was derived from an Rsa I digest of HFK cDNA amplified by PCR using the SJ1 and HE1 oligonucleotides. The cytoplasmic domain probe was a 140-bp fragment corresponding to nucleotides 1022-1162 of the human lymphoid sequence (Stamenkovic et al., 1989). It was derived by PCR amplification of HFF cDNA using the SJ2 and B7-2 oligonucleotides. Actin hybridization is shown for quantitative comparison of the total poly-A⁺ RNA loaded in each lane.

tion with chondroitinase (not shown). This form is therefore referred to as HFK HS/CS250. Thus, each cell population generated a sulfated form of CD44 with the high molecular mass forms expressing differing amounts of heparan sulfate, chondroitin sulfate, and other carbohydrate moieties.

Information about the nonglycosaminoglycan carbohydrates of GP230 and GP90 was obtained by lectin blotting and immunoblotting with carbohydrate-specific mAbs (Fig. 5). Purified GP90 and GP230 (purification of CD44 is discussed later under Fig. 10) were fractionated by SDS-PAGE, transferred to nitrocellulose, incubated with and without neuraminidase, and reacted with labeled lectins or carbohydrate-specific mAbs. Wheat germ agglutinin reacted strongly with both GP90 and GP230 before, but not after, digestion with neuraminidase, indicating the presence of extensive terminal sialylation. Removal of terminal sialic acid from HT1080 GP90 and QG56 and GP230 exposed penultimate galactose residues detected by binding of *Ricinus communis* agglutinin (gal-glcNAc preference) and peanut agglutinin (gal-galNac preference). The binding of the galactose-specific lectins was extensive on GP230 and weak on GP90. The binding of peanut agglutinin confirms the presence of ser/thr-linked, O-glycosides or mucin-like structures. After desialylation, GP230 bound the carbohydrate-specific mAb C6, that recognizes biantennary binary *N*-acetyllactosamine, a minimum polylactosamine structure (Nichols et al., 1987).

Identification of a Distinct Form of CD44 Core Protein, CD44E

We initially used pulse–chase labeling with [³⁵S]methionine to identify underglycosylated precursors of the fully glycosylated glycoproteins (Fig. 6). Two types of precursors were identified for the four characterized forms of CD44. HT1080 cells exhibit a single 65-kD precursor to GP90 (termed Pre-GP90). In contrast, four precursors to both QG56 GP230 and HFK HS/CS250 migrated with similar relative molecular masses of 90–135 kD (designated GPa-d or HSa-d, respectively), despite the differences in final glycosylation for these two forms.

Northern blot analysis provided further evidence that the core proteins for different CD44 forms may be similar, but not identical. A full-length cDNA probe to lymphoid CD44 detected at least three mRNA species in all cell populations examined (Fig. 7). However, epithelial cell types (i.e., HFKs and QG56 cells) exhibited major mRNA species between 1.0 and 1.6 kb longer (2.8, 3.6, and 6.7 kb) than what appeared to be their counterparts in fibroblasts (1.8, 2.4, and 5.1 kb). Upon longer exposure, the latter set of mRNA species was also apparent as minor species in the epithelial cells. It was possible that the larger set of mRNA species in these epithelial cells contained a larger coding region with unique sequences. Therefore, we amplified specific regions of cDNA
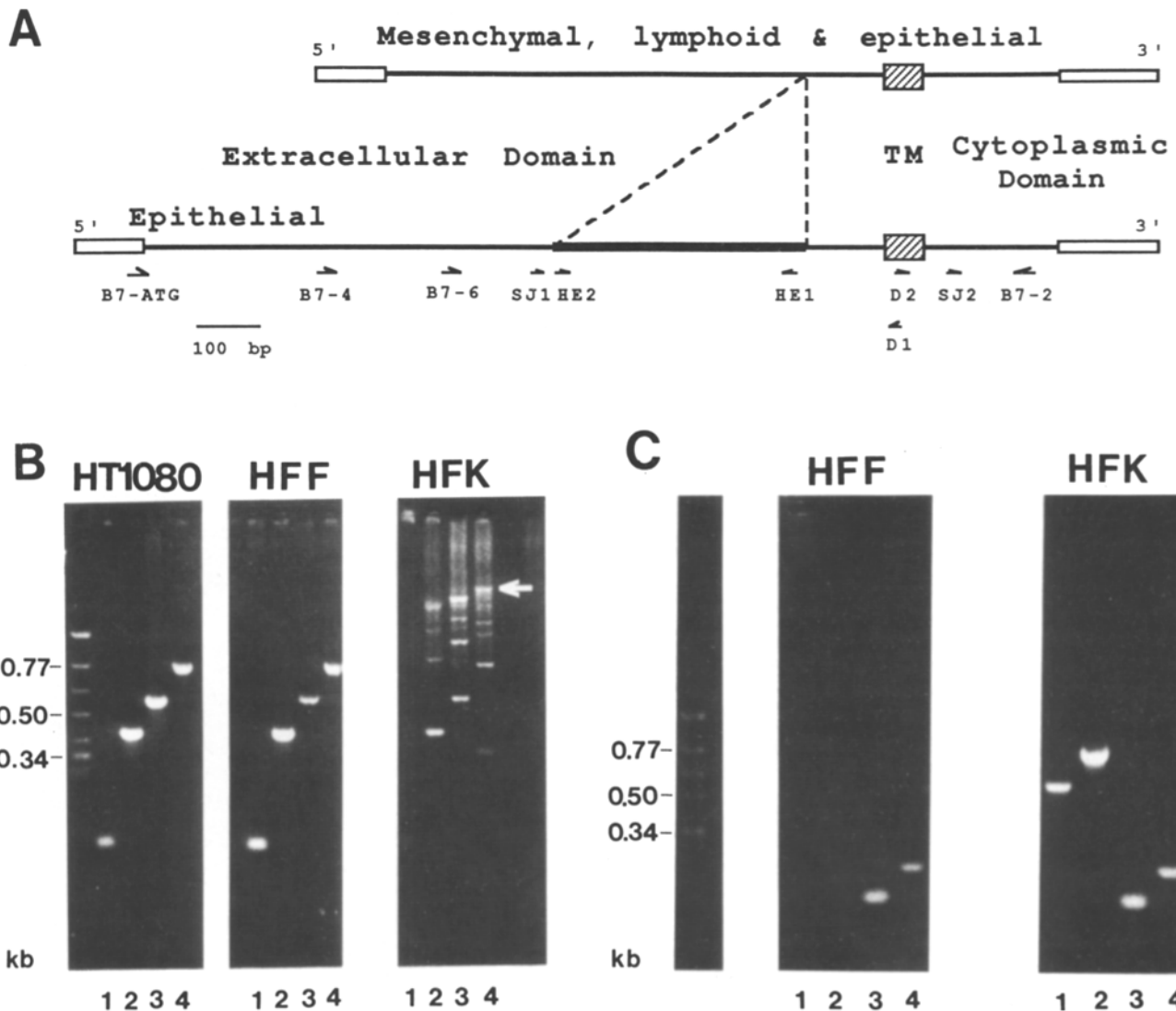
Figure 8. Analysis of CD44 mRNA species in epithelial and fibroblast cell populations by PCR amplification. (*A*) Linear map of CD44 and CD44E mRNAs with positions of oligonucleotides shown. The epithelial-associated coding region (*EACR, heavy bar*) and transmembrane region (*crosshatched bar*) are also shown. The untranslated regions (*open bars*) are taken from Stamenkovic et al. (1989). (*B*) PCR amplification of CD44 cDNA, reverse transcribed from HT1080, HFF and HFK poly-A+ RNA. For each cell type the oligonucleotides used were B7-2 and SJ2 (lanes *1*), SJ1 (lanes *2*), B7-6 (lanes *3*), or B7-4 (lanes *4*). Results identical to HFK lanes was obtained using QG56 cDNA. The cDNA band sequenced is denoted by the arrow. (*C*) PCR amplification of CD44E cDNA from HFK, but not HFF, cells using an EACR-specific oligonucleotide, HE2, and transmembrane oligonucleotide D1 (lanes *1*) or cytoplasmic oligonucleotide, B7-2 (lanes *2*) PCR amplification of EACR external sequences using B7-2 and SJ2 (lanes *3*) or D2 (lanes *4*) were included as controls.

from these cells using the PCR thermocycler with oligonucleotides specific for various regions of the lymphoid cDNA (Fig. 8 *A*). The reactions enlisted a common downstream or 3' oligonucleotide (i.e., B7-2 in the cytoplasmic protein domain, Fig. 8 *A*) coupled with various upstream (i.e., towards the extracellular protein domain) oligonucleotides (Fig. 8 *B*). Using the SJ2 oligonucleotide in conjunction with B7-2, only one cDNA of predicted length was amplified from both fibroblasts and epithelial cells (lanes *1*), suggesting that this region is similar in all transcripts that contain the cytoplasmic domain. When the upstream oligonucleotide B7-6, B7-4, or SJ1 was used in conjunction with B7-2, cDNAs amplified from HT1080 cells or HFFS were the same sizes as predicted for the lymphoid cDNA sequence. This suggested that the

fibroblast CD44 cDNA is identical to the lymphoid CD44 sequence. In contrast, amplification of cDNA from HFKs (Fig. 8 *B*) and QG56 cells (not shown) using the same combinations of oligonucleotides resulted in not only the predicted cDNA products from the lymphoid sequence, but also several larger cDNA products. As the upstream oligonucleotide was shifted more towards the 5' end, the epithelial-specific cDNA product obtained was uniformly larger. This indicated that the additional cDNA was inserted between the SJ1 and B7-2 oligonucleotides.

To identify the gel bands containing additional inserted cDNA sequences from HFKs and QG56 cells, individual DNA bands were excised from the gels, the cDNA isolated and used as templates in a subsequent, otherwise identical

```
agt gaa agg agc agc act tca gga ggt tac atc ttt tac acc ttt   709
 S   E   R   S   S   T   S   G   G   Y   I   F   Y   T   F     198
                                *
tct act gta cac ccc atc cca gac gaa gac agt ccc tgg atc acc   754
 S   T   V   H   P   I   P   D   E   D   S   P   W   I   T     213

gac agc aca gac aga atc cct gct acc aAT ATG GAC TCC AGT CAT   799
 D   S   T   D   R   I   P   A   T   N   M   D   S   S   H     228

AGT ACA ACG CTT CAG CCT ACT GCA AAT CCA AAC ACA GGT TTG GTG   844
 S   T   T   L   Q   P   T   A   N   P   N   T   G   L   V     243

GAA GAT TTG GAC AGG ACA GGA CCT CTT TCA ATG ACA ACG CAG CAG   889
 E   D   L   D   R   T   G   P   L   S   M   T   T   Q   Q     258

AGT AAT TCT CAG AGC TTC TCT ACA TCA CAT GAA GGC TTG GAA GAA   934
 S   N   S   Q   S   F   S   T   S   H   E   G   L   E   E     273

GAT AAA GAC CAT CCA ACA ACT TCT ACT CTG ACA TCA AGC AAT AGG   979
 D   K   D   H   P   T   T   S   T   L   T   S   S   N   R     288

AAT GAT GTC ACA GGT GGA AGA AGA GAC CCA AAT CAT TCT GAA GGC   1024
 N   D   V   T   G   G   R   R   D   P   N   H   S   E   G     303

TCA ACT ACT TTA CTG GAA GGT TAT ACC TCT CAT TAC CCA CAC ACG   1069
 S   T   T   L   L   E   G   Y   T   S   H   Y   P   H   T     318

AAG GAA AGC AGG ACC TTC ATC CCA GTG ACC TCA GCT AAG ACT GGG   1114
 K   E   S   R   T   F   I   P   V   T   S   A   K   T   G     333

TCC TTT GGA GTT ACT GCA GTT ACT GTT GGA GAT TCC AAC TCT AAT   1159
 S   F   G   V   T   A   V   T   V   G   D   S   N   S   N     348

GTC AAT CGT TCC TTA TCA Gga gac caa gac aca ttc cac ccc agt   1204
 V   N   R   S   L   S   G   D   Q   D   T   F   H   P   S     363
                        *                                *
ggg ggg tcc cat acc act cat gga tct gaa tca gat gga ca        1245
 G   G   S   H   T   T   H   G   S   E   S   D   G   :         376
```

Figure 9. cDNA and deduced protein sequences of the EACR from HFKs and QG56 cells. The EACR, spliced in immediately before nucleotide 783 of the human lymphoid sequence (Stamenkovic et al., 1989; the sequence numbers used are relative to this splice site), is composed of 396 nucleotides and shown in uppercase. The cDNA flanking the EACR (118 bp, 5' and 67 bp, 3') that was sequenced is shown in lower case. Potential O-glycosylation (*underline*), N-glycosylation (*bold underline*) and GAG attachment sites (∗) are indicated. The cDNA sequence obtained from both cell populations was identical and is available from EMBL/Genbank/DDBJ under accession number X55938.

PCR amplification. The largest PCR product (*arrow*, Fig. 8 *B*), resulted in much greater amplification of itself than the smaller lymphoid-associated product. This suggested that the largest product was a new epithelial cDNA product contaminated with minute amounts of the smaller lymphoid type cDNA (data not shown).

Both strands of the largest band from both HFKs and QG56 cells was sequenced directly using several of the oligonucleotides in Fig. 8 *A* as primers. The sequence was identical for both cell populations (Fig. 9), and contained an extra 396 bp between the 5' and 3' flanking lymphoid CD44 cDNA. We designated this extra cDNA the epithelial-associated coding region (EACR). Those transcripts that contain the EACR were designated as CD44E. The EACR was spliced into the CD44 mRNA before nucleotide 783 of the human sequence (Stamenkovic et al., 1989), the second nucleotide of the codon. This changed the arginine in the lymphoid form into an asparagine in the epithelial form. The EACR was unique in that it showed no significant homology to other sequences in the Genbank or EMBL databases (releases 63 and 23, respectively) and coded for an amino acid sequence composed of >30% serine and threonine residues. The cDNA sequence derived from both flanking regions agreed exactly with one of the reported human lymphoid CD44 sequences (Goldstein et al., 1989), and varied by only a single base from the other (Stamenkovic et al., 1989). In the latter a G was substituted for an A at nucleotide 831 of the lymphoid sequence (nucleotide 1227 of the epithelial

form), resulting in a glycine substitution for a glutamine in the amino acid sequence.

To determine which of the different CD44 mRNA transcripts contained the EACR sequence, Northern blots were probed with a 372-bp fragment contained within the EACR sequence (Fig. 7, *EACR*). All of the major mRNA species identified by the full-length lymphoid probe in QG56 cells and HFKs hybridized with the EACR-specific probe. In contrast, none of the major mRNA species in HT1080 cells or HFFs hybridized with the EACR probe. Using a cytoplasmic domain cDNA probe, we detected in the epithelial cells not only three major CD44E mRNA species, but also the three major lymphoid-associated CD44 transcripts (Fig. 7, *CYTO*). In addition, all four cell populations express a 1.3-kb transcript similar in size to the cloned CD44 sequence reported by Stamenkovic et al. (1989).
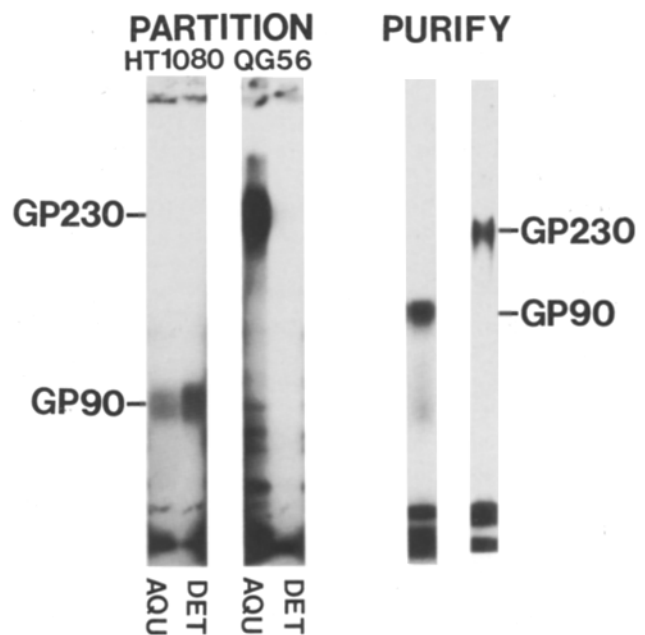

Figure 10. Differential extraction, phase partitioning with Triton X-114 detergent and purification of HT1080 GP90 and QG56 GP230. HT1080 and QG56 cells were ^{125}I-surface labeled, and sequentially extracted as previously described (Carter et al., 1982) first, with 2 M urea in 1 M NaCl to remove peripherally cell surface components, and second, with Triton X-114 (*PARTITION*) or X-100 (*PURIFY*) detergent at 4°C to solubilize membrane components followed by phase partitioning at 37°C for the Triton X-114 fractions (Bordier, 1981). The temperature increase results in formation of a detergent phase that contains membrane components with dominant hydrophobic character and an aqueous phase with dominant hydrophilic character. Each Triton X-114-derived phase was immune precipitated with anti-CD44 mAb (P1G12) and the antigens fractionated by SDS-PAGE followed by autoradiography. Unlabeled Triton X-100 extracts were used to affinity purify both HT1080 GP90 and QG56 GP230 by chromatography on WGA-Sepharose to purify all sialyated membrane glycoproteins followed by chromatography of the purified glycoproteins on P1G12-Sepharose. The bound and eluted GP90 and GP230 were labeled with ^{125}I Iodobeads and fractionated on SDS-PAGE (GP90 on 10%, GP230 on 8%) followed by autoradiography. Neither protein stained appreciably with Coomassie blue, possibly due to extensive glycosylation.

Intrinsic Membrane Character of CD44

The GP90 form of CD44 has been previously characterized as a transmembrane glycoprotein (Carter and Wayner, 1988) and a corresponding hydrophobic transmembrane sequence has been identified in the amino acid sequence of the core protein (Idzerda et al., 1989). No such information was available for the related higher molecular mass forms of CD44. We have used PCR amplification of CD44E cDNA from HFKs (Fig. 8 C) and QG56 cells (not shown) to demonstrate that at least one CD44E transcript includes very similar coding sequences for the transmembrane and cytoplasmic domains. Moreover, the 2.8- and 3.6-kb CD44E transcripts comigrate with a subset of the CD44 transcripts detected with a probe from the cytoplasmic domain (Fig. 7, *EACR* and *CYTO*). This suggested that these CD44E transcripts express cytoplasmic, and presumably hydrophobic transmembrane, domains similar to those in CD44. To analyze the polymorphic forms of CD44, we surface labeled HT1080 and QG56 cells with ^{125}I and subjected them to a sequential extraction as described in the legend to Fig. 10. Immunoprecipitation of the labeled cell extracts or immunoblotting of the unlabeled cell extracts with anti-CD44 mAbs detected the CD44 forms primarily in the Triton X-114 extract. As seen in Fig. 10 (*Partition*), temperature-induced phase partitioning of the Triton X-114 extract into detergent and aqueous phases further distinguished HT1080 GP90 and QG56 GP230; GP230 was recovered in the aqueous phase and as previously reported, GP90 was recovered in the detergent phase. In results not shown HT1080 CS180 and HFK HS/CS250 were recovered primarily in the Triton X-114 aqueous phase. Thus, all the CD44 forms required detergent for solubilization. However, the additional glycosylation and other possible modifications altered the partitioning character of the higher molecular mass forms.

We further establish that the larger polymorphic forms are intrinsic membrane glycoproteins with a hydrophobic domain that intercalates into the lipid bilayer of the plasma membrane. Unlabeled QG56 GP230 (and HT1080 GP90) was purified from Triton X-100 detergent extracts by affinity chromatography, and the purified proteins were labeled and incorporated into liposomes. Autoradiographs of the purified, labeled GP230 (and GP90) detected one major band in each preparation after SDS-PAGE (Fig. 10, *PURIFY*). Liposomes were prepared in the presence or absence of purified GP230. Only those liposomes containing GP230 would bind to immobilized anti-CD44 mAbs, indicating that GP230 incorporated into the liposomes in a proper configuration to interact with the antibodies (Table I).

Discussion

Peptide and Carbohydrate Polymorphism in CD44

We have identified and characterized the CD44 antigen as a polymorphic intrinsic membrane proteoglycan/glycoprotein with alternate splicing of the core peptide as well as extraordinary diversity in both the type and quantity of carbohydrate moieties. All the forms of CD44 contained at least one type of sulfated carbohydrate unit as well as nonsulfated carbohydrates. Previous reports relating to the CD44 have focused on hematopoietic cells (Jalkanen et al., 1988; Goldstein et

Table I. Incorporation of Purified GP230 from QG56 Cells into Liposomes: GP230 Is an Intrinsic Membrane Glycoprotein

Adhesive surface	Liposomes + GP230	Liposomes − GP230
BSA	516	644
IgG	594	681
P3H9	7,612	58

GP230 was purified as described in Fig. 10 and then incorporated into liposomes prepared from [3H]phosphatidylcholine, unlabeled egg yolk phosphatidylcholine, and octylglucoside detergent by dialysis (Mimms et al., 1981; Pytela et al., 1985). The labeled liposomes prepared with and without GP230 were incubated with plastic surfaces that had been coated with anti-CD44 antibody (P3H9) or control IgG and BSA (Carter and Wayner, 1988). The bound liposomes were solubilized and quantitated in a scintillation counter.

al., 1989; Stamenkovic et al., 1989; Pals et al., 1989) and have identified chondroitin sulfate as a carbohydrate moiety of the relatively minor higher molecular mass forms. The present study has established that epithelial cells, in particular, express the higher molecular mass forms as the major species of CD44. Glycosylation of these forms can be unique for each cell type with heparan sulfate being the major carbohydrate derivative of epidermal cells. Glycosylation of the related core proteins may reflect major differences in the glycosylation pathways of the individual cells and/or functional differences for the polymorphic forms (Paulson, 1989). Others have reported cell-type specific polymorphic forms of cell surface proteoglycans such as syndecan (Sanderson and Bernfield, 1988). Based on cDNA sequences, however, CD44E and human syndecan (Mali et al., 1990) have unrelated core proteins.

Lymphoid and mesenchymal cells express three major mRNA species for CD44, all presumably coding for a 37-kD core protein (Idzerda et al., 1989; Stamenkovic et al., 1989). In contrast, HFKs and QG56 cells express mRNA species for an alternate core protein, CD44E, in addition to the three mRNAs for CD44. The CD44E mRNAs code for a core protein with an additional 14 kD inserted in the extracellular domain that we have designated the EACR. The DNA sequence flanking the EACR was identical to that previously reported for the lymphoid core protein. Thus, the CD44E mRNA species probably arise from alternative splicing of a common gene rather than the use of an alternate CD44E gene.

The extra 132 amino acids in the CD44E core protein are spliced in the region near the COOH-terminal portion of the extracellular domain. The sequence flanking the splice site contains four repeats of the SGXG motif (Stamenkovic et al., 1989; Goldstein et al., 1989), a presumptive attachment site for chondroitin sulfate chains (Bourdon et al., 1987). Interestingly, the EACR is inserted in between these SGXG repeats and contains only one additional SG dipeptide at the COOH terminus. This extra repeat is a poor candidate for a chondroitin sulfate attachment site, though, because it contains neither upstream acidic residues nor the complete SGXG motif described by Bourdon et al. (1987). However, five potential O-glycosylation sites composed of at least three consecutive serine or threonine residues (Tomita et al., 1978) are found in the EACR sequence, adding to three similar sites in the flanking regions. Thus, QG56 cells may use these sites for the attachment of O-linked carbohydrate chains to GP250. The presence of heparan sulfate chains on the HFK

HS/CS250 and the QG56 "HEP" released form, but not on other forms of CD44, suggests that the EACR sequence may also contain as yet undescribed heparan sulfate attachment sites. Alternatively, the presence of the EACR may induce a new protein conformation exposing distant and previously buried heparan sulfate attachment sites in the flanking regions.

The expression of CD44E mRNAs by HFKs and QG56 cells correlated with the observation by pulse/chase labeling of higher mass precursors. QG56 GP230 and HFK HS/CS250 appear to originate from precursors that were distinct from the 65-kD HT1080 GP90 precursor. The QG56 and HFK precursors were similar, if not identical, despite the differences in final glycosylation. Nevertheless, we cannot conclude that the GP230 and HS/CS250 forms contain the CD44E core protein since mRNA species for both core proteins are expressed by these cells.

HT1080 GP90 contains a phosphorylated cytoplasmic domain that is linked to the glycosylated extracellular domain by a transmembrane hydrophobic sequence of amino acids based on both biochemical (Carter and Wayner, 1988) and cDNA sequence data (Idzerda et al., 1989; Stamenkovic et al., 1989). Similarly, the QG56 GP230 and HFK HS/CS250 both required detergent (but not high salt or urea) for solubilization and could be incorporated into liposomes, indicating the presence of a dominant hydrophobic domain. Moreover, at least two of the CD44E mRNA species codes for transmembrane and cytoplasmic domains very similar, if not identical, to the CD44 core proteins previously reported in lymphoid cells. This observation suggests that the CD44E core protein expressed by these mRNA species is not part of a form destined for secretion and may be used in the HFK HS/CS250 and QG56 GP230 forms. In addition, these CD44E core proteins would be expected to have similar associations with intracellular components, unlike the core protein reported by Goldstein et al. (1989) that lacks most of the cytoplasmic domain.

Functional Role for CD44 in Cell Adhesion

Filopodia have been suggested to be directly involved in initial cell–cell and cell–substratum contacts (Albrecht-Buehler et al., 1976). CD44 localizes in filopodia that contact with the substrate and in cell–cell contact sites suggesting that CD44 may form weak initial cell interactions. CD44 does not appear to be a cell-substrate adhesion receptor for collagen, fibronectin, or laminin (Wayner and Carter, 1987; Carter and Wayner, 1988). This conclusion is based on the absence of CD44 in focal adhesions, a major site of stable cell adhesion, and the inability of anti-CD44 antibodies to inhibit cell adhesion to ECM ligands. However, CD44 is an intrinsic membrane proteoglycan, a group of macromolecules that have been proposed as mediators of cell–substrate adhesion (Cole et al., 1985; Koda and Bernfield, 1984; Laterra et al., 1983; Woods et al., 1984, 1985; LeBaron et al., 1988; Ruoslahti, 1988). We have recently identified an alternate adhesion structure, termed a stable anchoring contact (SAC; Carter et al., 1990b). The possibility that CD44 interacts with SACs and/or other adhesion structures is currently under investigation. As a possible mechanism, CD44 expresses a partial sequence homology with mouse link protein (Idzerda et al., 1989; Goldstein et al., 1989; Stamen-

kovic et al., 1989). Link protein stabilizes the interaction of hyaluronate with the core protein of proteoglycans in cartilage (Neame et al., 1987). Recently, hyaluronate was shown to be a possible ligand for CD44 (Miyake et al., 1990b; Aruffo et al., 1990).

The possible role of CD44 in mediating cell adhesion may make use of the variability in carbohydrate content. The presence of major amounts of glycosaminoglycan on the core protein of HS/CS250 may modulate cell–cell interactions directly or via components of the extracellular matrix. Normal differentiation, stratification as well as abnormal lesions of epithelium have been reported to correlate with the differential expression of carbohydrate structures in skin (Dabelstein et al., 1986). CD44 is one of the major glycoconjugates of differentiating epidermis and exhibits an extraordinary sensitivity to regulation in carbohydrate content. This suggests a possible regulatory role for glycosylation in the function of CD44 in epidermis. Whether or not CD44 is primarily an adhesive glycoconjugate, its extensive polymorphism in both protein and carbohydrate dictate that care is required to identify the CD44 isoform responsible for a particular event when examining functional corollaries.

References

Albrecht-Buehler, G. 1976. The function of filopodia in spreading 3T3 mouse fibroblasts. In Cell Motility. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY. 247–264.

Alho, A. M., and C. B. Underhill. 1989. The hyaluronate receptor is preferentially expressed on proliferating epithelial cells. J. Cell Biol. 108:1557–1565.

Aruffo, A., I. Stamenkovic, M. Melnick, C. B. Underhill, and B. Seed. 1990. CD44 is the principle cell surface receptor for hyaluronate. Cell. 61:1303–1313.

Badley, J. E., G. A., Bishop, T. St. John, and J. A. Fralinger. 1988. A simple rapid method for the purification of poly A⁺ RNA. BioTechniques. 6:114–116.

Belitsos, P. C., J. E. K. Hildreth, and J. T. August. 1990. Homotypic cell aggregation induced by anti-CD44 (Pgp-1) monoclonal antibodies and related to CD44 (Pgp-1) expression. J. Immunol. 144:1661–70.

Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. J. Biol. Chem. 256:1604–7.

Bourdon, M.A., T. Krusins, S. Campbell, N. B. Schwartz, and E. Ruoslahti. 1987. Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. Proc. Natl. Acad. Sci. USA. 84:3194–3198.

Boyce, S. T., and R. G. Ham. 1985. Cultivation, frozen storage, and clonal growth of normal human epidermal keratinocytes in serum free media. J. Tissue Cult. Methods 9:83–93.

Burridge, K., K. Fath, T. Kelly, G. Nuckolls, and C. Turner. 1988. Focal adhesions: Transmembrane junctions between the extracellular matrix and the cytoskeleton. Annu. Rev. Cell Biol. 4:487–525.

Carter, W. G. 1982. Transformation-dependent alterations in glycoproteins of the extracellular matrix of human fibroblasts. J. Biol. Chem. 257:13805–13815.

Carter, W. G., and E. A. Wayner. 1988. Characterization of a collagen-binding, phosphorylated, transmembrane glycoprotein expressed in nucleated human cells. J. Biol. Chem. 263:4193–4201.

Carter, W. G., E. A. Wayner, T. S. Bouchard, and P. Kaur. 1990a. The role of integrins α2β1 and α3β1 in cell-cell and cell-substrate adhesion of human epidermal cells. J. Cell Biol. 110:1387–404.

Carter, W. G., P. Kaur, S. G. Gil, and E. A. Wayner. 1990b. Distinct functions for integrins α3β1 in focal adhesions and α6β4/bullous pemphigoid antigen in a new stable anchoring contact (SAC) of keratinocytes: relationship to hemidesmosomes. J. Cell Biol. 111:3141–3154.

Cole, G. J., D. Schubert, and L. Glaser. 1985. Cell-substratum adhesion in chick neural retina depends upon protein-heparan sulfate interactions. J. Cell Biol. 100:1192–1199.

Dabelsteen, E., K. Holbrook, H. Clausen, and S. Hakomori. 1986. Cell surface carbohydrate changes during embryonic and fetal skin development. *J. Invest. Dermatol.* 87:81-85.

Dalchau, R., J. Kirkley, and J. W. Fabre. 1980. Monoclonal antibody to a human brain-granulocyte-T-lymphocyte antigen probably homologous to the W3/13 antigen of the rat. *Eur. J. Immunol.* 10:745-749.

David, G., and H. Van den Berghe. 1985. Heparan sulfate-chondroitin sulfate hybrid proteoglycan of the cell surface and basement membrane of mouse mammary epithelial cells. *J. Biol. Chem.* 260:11067-11074.

Ennis, P. D., J. Zemmour, R. D. Salter, and P. Parham. 1990. Rapid cloning of HLA-A, B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. *Proc. Natl. Acad. Sci. USA.* 87:2833-2837.

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

Flanagan, B. F., R. Dalchau, A. K. Allen, A. S. Daar, and J. W. Fabre. 1989. Chemical composition and tissue distribution of the human CDw44 glycoprotein. *Immunology.* 67:167-175.

Gallatin, W. M., E. A. Wayner, P. A. Hoffman, T. St. John, E. C. Butcher, and W. G. Carter. 1989. Structural homology between lymphocyte receptors for high endothelium and class III extracellular matrix receptor. *Proc. Natl. Acad. Sci. USA.* 86:4654-4658.

Goldstein, L. A., D. F. H. Zhou, L. J. Picker, C. N. Minty, R. B. Bargatze, J. F. Ding, and E. C. Butcher. 1989. A human lymphocyte homing receptor, the hermes antigen, is related to cartilage proteoglycan core and link proteins. *Cell.* 56:1063-1072.

Grimm, E. A., A. Mazumder, H. Z. Zhang, and S. A. Rosenberg. 1982. Lymphokine activated killer cell phenomenon. Lysis of natural killer-resistant fresh solid tumor cells by interleukin 2-activated autologous human peripheral blood lymphocytes. *J. Exp. Med.* 155:1923-1941.

Haynes, B. F., E. A. Harden, M. J. Telen, M. E. Hemler, J. L. Strominger, T. J. Palker, R. M. Scearce, and G. S. Eisenbarth. 1983. Differentiation of human T lymphocytes. I. Acquisition of a novel human cell surface protein (p80) during normal intrathymic T cell maturation. *J. Immunol.* 131:1195-1200.

Haynes, B. F., M. J. Telen, L. P. Hale, and S. M. Denning. 1989. CD44: A molecule involved in leukocyte adherence and T-cell activation. *Immunol. Today.* 10:423-428.

Huet, S., H. Groux, B. Caillou, H. Valentin, A. Prieur, and A. Bernard. 1989. CD44 contributes to T cell activation. *J. Immunol.* 143:798-801.

Idzerda, R. L., W. G. Carter, C. Nottenburg, E. A. Wayner, W. M. Gallatin, and T. St. John. 1989. Isolation and DNA sequence of a cDNA clone encoding a lymphocyte adhesion receptor for high endothelium. *Proc. Natl. Acad. Sci. USA.* 86:4659-4663.

Isacke, C. M., C. A. Sauvage, R. Hyman, J. Lesley, R. Schulte, and I. S. Trowbridge. 1986. Identification and characterization of the human Pgp-1 glycoprotein. *Immunogenetics.* 23:326-332.

Izzard, C. S., and L. R. Lochner. 1976. Cell-to-substrate contacts in living fibroblasts: an interference reflexion study with an evaluation of the technique. *J. Cell Sci.* 21:129-159.

Jalkanen, J., R. F. Bargatze, J. de los Toyos, and E. C. Butcher. 1987. Lymphocyte recognition of high endothelium: antibodies to distinct epitopes of an 85-95-kD glycoprotein antigen differentially inhibit lymphocyte binding to lymph node, mucosal, or synovial endothelial cells. *J. Cell Biol.* 105:983-990.

Jalkanen, S., M. Jalkanen, R. Bargatze, M. Tammi, and E. C. Butcher. 1988. Biochemical properties of glycoproteins involved in lymphocyte recognition of high endothelial venules in man. *J. Immunol.* 141:1615-1623.

Kalomiris, E. L., and L. Y. W. Bourguignon. 1988. Mouse T lymphoma cells contain a transmembrane glycoprotein (GP85) that binds ankyrin. *J. Cell Biol.* 106:319-327.

King, P. D., A. H. Batchelor, P. Lawlor, and D. R. Katz. 1990. The role of CD44, CD45, CD45RO, CD46 and CD55 as potential anti-adhesion molecules involved in the binding of human tonsillar T cells to phorbol 12-myristate 13-acetate-differentiated U-937 cells. *Eur. J. Immunol.* 20:363-368.

Koda, J. E., and M. Bernfield. 1984. Heparan sulfate proteoglycans from mouse mammary epithelial cells. *J. Biol. Chem.* 259:11763-11770.

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

Laterra, J., J. E. Silbert, and L. A. Culp. 1983. Cell surface heparan sulfate mediates some adhesive responses to glycosaminoglycan-binding matrices, including fibronectin. *J. Cell Biol.* 96:112-123.

LeBaron, R. G., J. D. Esko, A. Woods, S. Johansson, and M. Hook. 1988. Adhesion of glycosaminoglycan-deficient chinese hamster ovary cell mutants to fibronectin substrata. *J. Cell Biol.* 106:945-952.

Lesley, J., R. Schulte, and R. Hyman. 1990. Binding of hyaluronic acid to lymphoid cell lines is inhibited by monoclonal antibodies against Pgp-1. *Exp. Cell Res.* 187:224-233.

LeTarte, M., S. Iturbe, and E. J. Quackenbush. 1985. A glycoprotein of molecular weight 85,000 on human cells of B-lineage: detection with a family of monoclonal antibodies. *Mol. Immunol.* 22:113-124.

Mali, M., P. Jaakkola, A. M. Arvilommi, and M. Jalkanen. 1990. Sequence of human syndecan indicates a novel gene family of integral membrane proteoglycans. *J. Biol. Chem.* 265:6884-6889.

Miyake, K., K. L. Medina, S. Hayashi, S. Ono, T. Hamaoka, and P. W. Kincade. 1990a. Monoclonal antibodies to Pgp-1/CD44 block lymphohemopoiesis in long-term bone marrow cultures. *J. Exp. Med.* 171:477-488.

Miyake, K., C. B. Underhill, J. Lesley, and P. W. Kincade. 1990b. Hyaluronate can function as a cell adhesion molecule and CD44 participates in hyaluronate recognition. *J. Exp. Med.* 172:69-75.

Mimms, L. T., G. Zampighi, Y. Nozaki, C. Tanford, and J. A. Reynolds. 1981. Phospholipid vesicle formation and transmembrane protein incorporation using octyl glucoside. *Biochemistry.* 20:833-840.

Neame, P. J., J. E. Christner, and J. R. Baker. 1987. Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-terminal globular domains have similar structures. *J. Biol. Chem.* 262:17768-17778.

Neyfakh, A. A., I. S. Tint, T. M. Svitkina, A. D. Bershadsky, V. I. Gelfand. 1983. Visualization of cellular focal contacts using a monoclonal antibody to 80 kd serum protein adsorbed on the substratum. *Exp. Cell Res.* 149:387-96.

Nichols, E. J., B. A. Fenderson, W. G. Carter, and S. Hakomori. 1986. Domain-specific distribution of carbohydrates in human fibronectins and the transformation-dependent translocation of branched type 2 chain defined by monoclonal antibody C6*. *J. Biol. Chem.* 261:11295-301.

Omary, M. B., I. S. Trowbridge, M. Letarte, M. F. Kagnoff, and C. M. Isacke. 1988. Structural heterogeneity of human Pgp-1 and its relationship with p85. *Immunogenetics.* 27:460-464.

Pals, S. T., F. Hogervorst, G. D. Keizer, T. Thepen, E. Horst, and C. C. Figdor. 1989. Identification of a widely distributed 90-kDa glycoprotein that is homologous to the hermes-1 human lymphocyte homing receptor. *J. Immunol.* 143:851-857.

Parikh, H., S. March, and P. Cuatrecasas. 1974. Topics in the methods of substitution reactions with agarose. *Methods Enzymol.* 34:77-101.

Paulson, J. C. 1989. Glycoproteins: What are the sugar chains for? *Trends Biol. Sci.* 14:272-276.

Picker, L. J., J. de los Toyos, M. J. Telen, B. F. Haynes, and E. C. Butcher. 1989. Monoclonal antibodies against the CD44 [In(Lu)-related p80], and Pgp-1 antigens in man recognize the hermes class of lymphocyte homing receptors. *J. Immunol.* 142:2046-2051.

Pytela, R., M. D. Pierschbacher, and E. Ruoslahti. 1985. Identification and isolation of a 140 kd cell surface glycoprotein with properties expected of a fibronectin receptor. *Cell.* 40:191-198.

Rice, R. H., and H. Green. 1979. Presence in human epidermal cells of a soluble protein precursor of the cross-linked envelope: activation of the cross-linking by calcium ions. *Cell.* 18:681-694.

Roman, J., R. LaChance, T. J. Broekelmann, C. J. Roberts, E. A. Wayner, W. G. Carter, and J. McDonald. 1989. The fibronectin receptor is organized by extracellular matrix fibronectin: Implication for oncogenic transformation and for cell recognition of fibronectin matrices. *J. Cell Biol.* 108:2529-2543.

Ruoslahti, E. 1988. Structure and biology of proteoglycans. *Annu. Rev. Cell Biol.* 4:229-255.

Sanderson, R. D., and M. Bernfield. 1988. Molecular polymorphism of a cell surface proteoglycan: distinct structures on simple and stratified epithelia. *Proc. Natl. Acad. Sci. USA.* 85:9562-9566.

Sanger, F., S. Miklen, and A. R. Coulsen. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.

Selden, R. F. 1989. Analysis of RNA by Northern hybridization. *In* Current Protocols in Molecular Biology. Greene Publishers, New York. 4.9.1-8.

Shimizu, Y., G. A. Van Seventer, R. Siraganian, L. Wahl, and S. Shaw. 1989. Dual role of the CD44 molecule in T cell adhesion and activation. *J. Immunol.* 143:2457-2463.

Shively, J. E., and H. E. Conrad. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry* 15:3932-3942.

St. John, T., J. Meyer, R. Idzerda, and W. M. Gallatin. 1990. Expression of CD44 confers a new adhesive phenotype on transfected cells. *Cell.* 60:45-52.

Stamenkovic, I., M. Amiot, J. M. Pesando, and B. Seed. 1989. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell.* 56:1057-1062.

Tautz, D., and M. Renz. 1983. An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. *Anal. Biochem.* 132:14-19.

Tomita, M., H. Furthmayr, and V. T. Marchesi. 1978. Primary structure of human erythrocyte glycophorin A: isolation and characterization of peptides and complete amino acid sequence. *Biochemistry.* 17:4756-70.

Undenfriend, S., S. Stein, P. Bohlen, W. Dairman, W. Leingruber, and M. Weigle. 1972. Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science (Wash. DC).* 178:871-872.

Underhill, C. B. 1989. The interaction of hyaluronate with the cell surface: the hyaluronate receptor and the core protein. *In* The Biology of Hyaluronan. John Wiley & Sons, Chichester, UK. 87-106.

Wayner, E. A., and W. G. Carter. 1987. Identification of multiple cell adhesion receptors for type VI collagen and fibronectin in human fibrosarcoma cells possessing unique α and common β subunits. *J. Cell Biol.* 105:1873-1884.

Wayner, E. A., W. G. Carter, R. S. Piotrowicz, and T. J. Kunicki. 1988. The function of multiple extracellular matrix receptors in mediating cell adhesion

220

to extracellular matrix: preparation of monoclonal antibodies to the fibronectin receptor that specifically inhibit cell adhesion to fibronectin and react with platelet glycoproteins Ic-IIa. *J. Cell Biol.* 107:1881–1891.

Winship, P. R. 1989. An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.* 17:1266.

Woods, A., J. R. Couchman, and M. Hook. 1985. Heparan sulfate proteoglycans of rat embryo fibroblasts. *J. Biol. Chem.* 260:10872–10879.

Woods, A., M. Hook, L. Kjellen, C. G. Smith, and D. A. Rees. 1984. Relationship of heparan sulfate proteoglycans to the cytoskeleton and extracellular matrix of cultured fibroblasts. *J. Cell Biol.* 99:1743–1753.