Biochemistry

 Volume 29, Number 43 October 30, 1990

Perspectives in Biochemistry

The Asialoglycoprotein Receptor: A Model for Endocytic Transport Receptors

Martin Spiess

Department of Biochemistry, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received May 4, 1990; Revised Manuscript Received June 22, 1990

There are two basic mechanisms for the uptake of substances into the cell. Small molecules such as amino acids, monosaccharides, and most ions are transported directly across the plasma membrane into the cytoplasm. Macromolecules like peptides and proteins are taken up by the process of receptor-mediated endocytosis (Goldstein et al., 1985); they are specifically recognized and bound with high affinity to receptors on the cell surface. These are then clustered into clathrin-coated patches of the plasma membrane called coated pits, which invaginate and pinch off as coated vesicles. The cytoplasmic clathrin coat is rapidly removed, and the vesicles fuse with an acidic endosomal compartment. Although inside the cell, the ligands are still separated from the cytoplasm by a membrane. In most cases, they are delivered to lysosomes and hydrolyzed before their degradation products are released into the cytoplasm.

The endosome is a pivotal sorting compartment: distinct vesicles depart from this compartment either toward prelysosomes and lysosomes (the degradative pathway) or back toward the plasma membrane (the recycling pathway). In polarized cells, which have separate apical and basolateral plasma membrane domains, there are recycling routes from endosomes to either one of them and transcytotic routes to the opposite surface. Several receptor classes can be distinguished by the pathways followed by receptor and ligand. Transcytotic receptors (e.g., the polymeric immunoglobulin [poly(Ig)] receptor; Mostov & Simister, 1985) remain bound to their ligands during their journey across the epithelial cell. Some hormone receptors (e.g., EGF receptor; Yarden & Ullrich, 1988; Schlessinger, 1988) are transported with their ligands to lysosomes where both are degraded. Recycling receptors usually carry their ligands only to the endosomes where an acid-induced conformational change causes dissociation of the complex; the released ligands take the degradative route, while the receptors are sorted into recycling vesicles that bring them back to the plasma membrane for reuse. They have the characteristics of efficient, high-rate uptake systems. Accordingly, the ligands are frequently either nutrients for the cell [e.g., cholesterol-containing low-density lipoproteins (LDL) and iron-containing transferrin] or macromolecules to be cleared from the circulation (e.g., desialylated glycoproteins and α_2-macroglobulin–protease complexes).

The best studied members of the class of recycling transport receptors are the LDL receptor (Goldstein et al., 1985), the transferrin receptor (May & Cuatrecasas, 1985), and the asialoglycoprotein (ASGP) receptor, the subject of this paper. Although these receptors use the same general pathways, their molecular structures have little in common. This paper focuses on the ASGP receptor for which recently obtained structural information is just beginning to reveal some of the mechanisms underlying receptor trafficking.

A Mammalian Hepatic Lectin

The ASGP receptor was originally discovered during studies on the metabolism of serum glycoproteins: upon removal of the terminal sialic acid residues on their N-linked oligosaccharides, they are rapidly cleared from the circulation and degraded in liver cells [reviewed by Ashwell and Harford (1982)]. It was discovered that terminal galactose residues in asialoglycoproteins (ASGPs) are recognized by a receptor whose binding site is specific for D-galactose, D-*N*-acetylgalactosamine, and related galactosides [reviewed by Schwartz (1984b)]. Accordingly, the ASGP receptor has been alternatively called "hepatic lectin" and "galactose (or galactose/*N*-acetylgalactosamine) receptor". The existence of a pool of circulating ASGPs (with potential diagnostic value) in patients with diseases of the liver (Sawamura et al., 1984) suggests a physiological function for the hepatic receptor in serum glycoprotein turnover in concert with a desialylating activity that acts stochastically on glycoproteins in the circulation. However, there is no direct evidence for this general function, and it is possible that a more specific physiological ligand exists [e.g., IgA in humans (Daniels et al., 1989) and lipoprotein A (Soutar, 1989)]. Consistent with the general function proposed above, the serum of birds contains considerable amounts of galactose-terminal glycoproteins (Regoeczi et al., 1975), and only upon degalactosylation are they efficiently cleared by a hepatic receptor specific for terminal *N*-acetylglucosamine residues. The chicken hepatic lectin has been characterized and was found to be partly homologous with the mammalian

0006-2960/90/0429-10009$02.50/0

ASGP receptor (Lunney & Ashwell, 1976; Drickamer, 1981; see also below).

Beyond the specificity for terminal galactose residues, binding to the ASGP receptor depends strongly on the structure of the oligosaccharide: mono-, bi-, tri-, and tetraantennary galactose-terminal oligosaccharides bind with increasing affinities, with dissociation constants of 10^{-3}, 10^{-6}, 5×10^{-9}, and 10^{-9} M, respectively (Lee et al., 1983). This suggests a close arrangement of several galactose-binding sites within the span of a ligand oligosaccharide. There is also evidence that ligands containing several oligosaccharides engage in multiple interactions (Hardy et al., 1985), which must be accounted for by the spatial organization of the receptor. Ligand binding has an absolute requirement for Ca^{2+} (optimal concentration of 0.1–2 mM; Weigel, 1980; Van Lenten & Ashwell, 1972) and is effective only above pH 6.5 (Schwartz et al., 1983), providing the molecular basis for ligand dissociation in acidic endosomes.

The receptor is expressed exclusively by parenchymal hepatocytes (Hubbard et al., 1979; Hubbard & Stukenbrok, 1979) which contain 100 000–500 000 binding sites per cell. Surface receptors are randomly distributed over the sinusoidal (basolateral) plasma membrane domain facing the capillaries and are enriched in coated pits, but they are essentially absent from the apical membrane facing the bile canaliculi (Wall & Hubbard, 1981; Matsuura et al., 1982; Geuze et al., 1982; 1983b). In the human hepatoma cell line HepG2, it was found that at high ligand concentration binding occurs within approximately 1 min, that internalization occurs with a mean time of 2 min, and that an additional 4 min is required for ligand dissociation and reappearance of the receptor at the cell surface, resulting in a minimum cycle time of 7 min (Schwartz et al., 1982). The majority of the dissociated ASGPs reaches lysosomes within 5–15 min. With a mean liftime of 30 h, a single receptor can thus internalize up to 250 ligand molecules. Similar parameters have been determined for other recycling receptors such as those for LDL and transferrin (Breitfeld et al., 1985).

The intracellular pathways of receptor and ligand have been delineated in detail by electron microscopy [Geuze et al. (1983a,b, 1984), Mueller and Hubbard (1986), and others reviewed by Schwartz (1984b) and Breitfeld et al. (1985)]. ASGPs and their receptor have been found in coated pits and coated and uncoated vesicles together with the mannose 6-phosphate, transferrin, and poly(Ig) receptors (Geuze et al., 1984; Stoorvogel et al., 1989). In endosomes (also called "compartment for the uncoupling of receptor and ligand"), ASGPs are dissociated from their receptors in the acidic environment produced by an endosomal H^+-translocating ATPase (pH 5–6; Mellman et al., 1986). In this compartment, the ASGP receptor is segregated from the poly(Ig) receptor, which is transcytosed (Geuze et al., 1984). The ASGP receptor is recycled to the plasma membrane by vesicles budding from tubular extensions of endosomes.

Like other recycling receptors (e.g., those for LDL and transferrin; Goldstein et al., 1985; Watts, 1985), the ASGP receptor is endocytosed and recycled constitutively. In the presence or absence of ligand, a considerable portion of the receptor is always found intracellularly. Treatment with lysosomotropic agents (such as ammonium chloride, chloroquine, and monensin) neutralizes the endosomal pH and blocks receptor recycling but not internalization. This causes a rapid and reversible loss of cell surface binding sites even in the absence of ligand (Tolleshaug & Berg, 1979; Tycko et al., 1983; Schwartz et al., 1984). Constitutive endocytosis and reappearance of surface-labeled receptors have been shown directly in transfected fibroblasts expressing functional ASGP receptors (Geffen et al., 1989). However, ligand binding does increase the rate of endocytosis, resulting in a rapid and transient reduction of cell surface binding sites upon ligand addition (Ciechanover et al., 1983; Schwartz et al., 1984).

Two Genes, Two Proteins, One Receptor

The ASGP receptor proteins can be conveniently purified by affinity chromatography on immobilized ASGPs or galactose, advantage being taken of the pH and Ca^{2+} dependence of galactose binding for adsorption and elution. By this procedure, two proteins of 40 and 48 kDa in a ratio of approximately 2:1 were isolated from rabbit liver (Kawasaki & Ashwell, 1976). From rat liver three proteins (called rat hepatic lectins RHL-1, -2, and -3) of approximately 42, 49, and 54 kDa in a ratio of 8:1:1 (Drickamer et al., 1984) were isolated, and similarly three proteins from mouse (Hong et al., 1988). Only the receptor from human liver and HepG2 cells was purified as an apparently single protein of 46 kDa (Baenziger & Maynard, 1980; Schwartz & Rup, 1983).

The ASGP receptor proteins are integral membrane proteins that require detergents for solubilization. They span the lipid bilayer once and expose a relatively small amino-terminal segment to the cytoplasm and a large carboxy-terminal portion to the exoplasmic milieu (Chiaccia & Drickamer, 1984). During biosynthesis they are cotranslationally inserted into the endoplasmic reticulum and modified by the addition of high-mannose-type oligosaccharides at two or three sites for N-linked glycosylation (Asn-X-Ser/Thr). The oligosaccharides are converted to the complex type within 45–60 min (in HepG2 cells; Schwartz & Rup, 1983; Bischoff & Lodish, 1987) as the protein is transported through the Golgi apparatus to the plasma membrane. Inhibition of glycosylation by tunicamycin (which blocks the formation of the dolichyl oligosaccharide precursor) or of maturation by swainsonine (which inhibits Golgi mannosidase II) has little effect on the half-life of the receptor and does not affect ligand binding and internalization (Breitfeld et al., 1984; Hsueh et al., 1986).

The different proteins purified are immunologically cross-reactive within and between species, reflecting a close relationship. Sequencing of the rat proteins revealed that the two minor receptor proteins RHL-2 and RHL-3, in fact, have an identical amino acid sequence which is different but highly homologous to that of the major form RHL-1 (Drickamer et al., 1984). This suggested the existence of two genes encoding ASGP receptors in rat, a notion that was finally confirmed by the cloning and sequencing of the corresponding two cDNAs (Holland et al., 1984; McPhaul & Berg, 1987; Halberg et al., 1987). The two minor rat proteins RHL-2 and RHL-3 differ only in their carbohydrate structures as was shown by endoglycosidase digestion and lectin binding assays (Halberg et al., 1987).

The existence of two genes in the human system was only discovered by cDNA cloning (Spiess et al., 1985; Spiess & Lodish, 1985). Oligopeptides corresponding to segments that differ in the two human receptors were then synthesized and used to raise antisera. With the help of these specific antisera, it was possible to demonstrate the expression of two ASGP receptor proteins in human HepG2 cells (Bischoff & Lodish, 1987). Although their mRNAs are detectable in approximately equal amounts (Spiess & Lodish, 1985), the human proteins (called H1 and H2) are translated in a ratio of 5–6:1 to give major and minor form as in other species.

Southern analyses performed on rat and mouse DNA suggest that there are only two genes encoding ASGP receptors

(Leung et al., 1985; McPhaul & Berg, 1987; Sanford et al., 1988). These two genes were shown to be closely linked on chromosome 11 of the mouse (Sanford et al., 1988). Since all three rat receptor species appear with similar time courses in development (Petell & Doyle, 1985), it has been suggested that the close linkage of the genes may enable them to be coordinately regulated by common cis elements.

cDNA and deduced protein sequences of the two human and the two rat ASGP receptors are highly homologous: 39% of the amino acids are identical in all four sequences. The two major species, H1 and RHL-1, have 80% identical amino acids and are more closely related to each other than to H2 (58% and 56%) and RHL-2/3 (52% and 50%, respectively). Likewise, H2 and RHL-2/3 are more homologous to each other (62%) than to H1 and RHL-1. Clearly, the two receptor genes must have been generated by gene duplication before primates and rodents diverged in evolution, at least 65 million years ago. Since then, the gene for the major receptor form has been more highly conserved than the gene for the minor form.

The homology extends over the entire coding sequence except for an 18-codon insertion in the minor sequence that does not have a counterpart in the major one. Leung et al. (1985) have cloned and sequenced the gene of RHL-1. The insertion point at codon 24 in the cytoplasmic domain of the protein coincides exactly with the position of intron 2, suggesting that the additional exon sequence in gene 2 is due to an incorrect splicing event in either one of the genes. This event must have happened before the separation of primates and rodents. Furthermore, approximately 10% of the mRNAs for H2 (designated H2A) are alternatively spliced at the position of intron 3, i.e., at the exoplasmic end of the membrane-spanning segment. H2A mRNAs contain a 5-codon insertion that corresponds to the 5′-end of intron 3 (Lederkremer and Lodish, personal communication). The possible functions of these insertion sequences are not known.

The existence and conservation of two homologous galactose-binding proteins in mammalian hepatocytes raised the question of whether they function as separate, independent receptors or as constituents of a single multicomponent receptor. Cross-linking of the rat receptors produced separate homooligomers of RHL-1 and RHL-2/3, up to hexamers in detergent solution and up to trimers in membranes (Halberg et al., 1987). However, the human receptors in HepG2 cells were cross-linked to mixed dimers and trimers containing H1 and H2 (Bischoff et al., 1988). When receptor degradation was induced in HepG2 cells by cross-linking with an antibody directed against H1 or H2, both polypeptides were observed to be internalized and degraded simultaneously, suggesting that the two receptor forms are assembled in heterooligomeric complexes (Bischoff et al., 1988). Similarly, specific antisera raised against carboxy-terminal peptides of either RHL-1 or RHL-2/3 both precipitated all three forms of the rat receptor from iodinated hepatocytes (Sawyer et al., 1988).

In a functional test, McPhaul and Berg (1986) found that HTC cells (a rat hepatoma cell line lacking ASGP receptors) acquired the ability to internalize and accumulate fluorescently labeled asialoorosomucoid only when transfected with the cDNAs of both rat proteins RHL-1 and RHL-2/3. the clearest evidence for the ASGP receptor being a heterooligomeric complex was provided by Shia and Lodish (1989), who analyzed the fate and properties of H1 and H2 when expressed separately or together in stably transfected fibroblasts. In the absence of H2, H1 is transported to the cell surface normally but does not bind ASGPs. In contrast, H2 expressed alone is rapidly degraded, still in its high-mannose glycosylated form, and does not reach the plasma membrane. In cells expressing both proteins, H2 is "rescued" to the cell surface, and ligand binding and uptake are observed as in HepG2 cells. Association of the two polypeptides is thus necessary for the transport of H2 to the plasma membrane and for high-affinity ligand binding. Heterooligomerization appears to create a productive arrangement of the subunit binding sites. All subunits of the complex are involved in ligand binding as judged by affinity labeling and cross-linking of ligand and receptor (Lee & Lee, 1986, 1987; Herzig & Weigel, 1989).

The target sizes of the high-affinity ligand binding sites of the ASGP receptor have been determined by radiation inactivation. The results, ~104, ~109, ~70, and ~140 kDa for rat, rabbit, human liver, and HepG2 membranes, respectively (Steer et al., 1981; Schwartz et al., 1984), reflect the oligomeric requirement for ligand binding; however, they do not allow an unambiguous determination of the number of subunits involved. Considering the structure of the smallest high-affinity ligand, a triantennary oligosaccharide, the minimal receptor unit appears to be a trimer. The simplest unit would therefore consist of two subunits 1 and one subunit 2. Alternatively, a complex of 5:1, which more closely represents the measured ratio of the major and minor receptor forms in hepatocytes, might present two high-affinity binding sites. In any case, the binding behavior of the receptor suggests that the high-affinity binding units are further arranged in (permanent or transient) clusters on the cell surface (Hardy et al., 1985). The molecular mass of the receptor complex in detergent solution of approximately 260 kDa (Andersen et al., 1982) is consistent with a hexamer, implying that this might be a preferred complex also in the membrane. The possibility that formation of higher complexes is induced by polyvalent ligands could provide a mechanism for the observed ligand-induced increase of internalization rate. The chicken hepatic lectin, composed of identical polypeptides, similarly forms hexamers (Loeb & Drickamer, 1987).

As mentioned above, expression of subunit H1 is a prerequisite for efficient transport of H2 to the plasma membrane in transfected fibroblasts (Shia & Lodish, 1989). Since H2 remains in the high-mannose glycosylated form in the absence of H1, it is probably retained in the ER. For various membrane proteins it has been shown that they need to be assembled in an oligomeric complex before exiting this organelle [e.g., Kreis and Lodish (1986) and Gething et al. (1986)]. While H1 alone does oligomerize, as judged by cross-linking experiments (Shia & Lodish, 1989), H2 alone might not. The retained H2 polypeptides are degraded within 60–90 min (Amara et al., 1989). Interestingly, the first step of degradation, the removal of approximately 8 kDa from the amino terminus, takes place in the ER, and further degradation appears to be nonlysosomal, as it is not affected by leupeptin (an inhibitor of lysosomal proteases) or by lysosomotropic agents.

Gene Organization

In many genes, exons correspond to functional domains within proteins (a striking example is the LDL receptor; Südhof et al., 1985). This appears also to be the case for the ASGP receptor genes (Figures 1A and 2). Most of the cytoplasmic domain is encoded by exon 2, the membrane-spanning segment by exon 3, and the carbohydrate-binding domain by exons 7–9. Exon 1 comprises the beginning of the 5′-noncoding sequence. The function of the middle exons 4–6 is not obvious. In RHL-1, an internal repeat with almost 30%

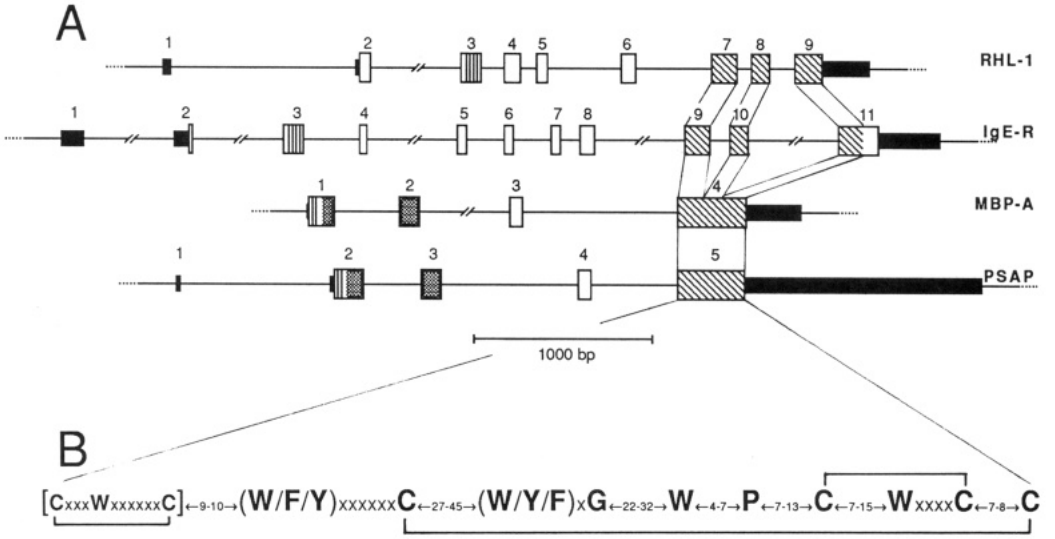

FIGURE 1: Gene structure of the ASGP receptor RHL-1 and related proteins with homologous carbohydrate-binding domains. (A) The gene organization of the rat ASGP receptor RHL-1 (Leung et al., 1985), the human IgE receptor (IgE-R; Suter et al., 1987), the rat mannose-binding protein A (MBP-A; Drickamer & McCreary, 1987), and the human pulmonary surfactant apoprotein (PSAP; White et al., 1985) is shown. Exons are indicated by boxes (narrow, filled ones for noncoding sequences). The hatched, stippled, and striped segments encode carbohydrate-binding domains, collagen-like sequences, and hydrophobic signal or transmembrane sequences, respectively. (B) The residues within the carbohydrate-binding domain that are conserved in 27 proteins [referenced in Drickamer (1988), Petersen (1988), Patthy (1988), and Thiel and Reid (1989)] are shown. The segment in brackets is conserved only in a subpopulation of 12 proteins including the ASGP receptor proteins. Variable residues are indicated by X, and longer nonconserved peptides are indicated by their lengths. Cysteines forming disulfide bonds are connected by lines.

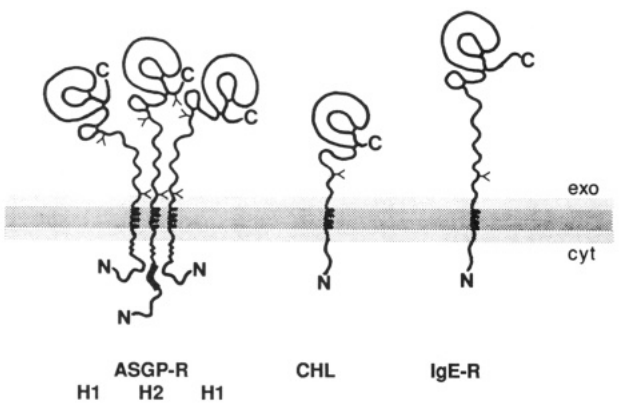

FIGURE 2: Schematic structure of the ASGP receptor. For simplicity, a trimer of two subunits H1 and one subunit H2 is drawn. For comparison, a single polypeptide of the chicken hepatic lectin (CHL) and of the IgE receptor (IgE-R) is depicted. The cytoplasmic (cyt) and exoplasmic (exo) sides of the membrane are indicated.

identical amino acids has been noticed between residues 122–144, corresponding almost exactly to exon 6, and residues 51–77, which encompass regions of exons 3 and 4 and are interrupted by intron 3 (Drickamer et al., 1984). This repeat is also detectable in H1, whereas no significant internal homology is found in H2 and RHL-2/3.

The chicken hepatic lectin, the (presumed) avian analogue of the mammalian ASGP receptors, has a similar overall structure, yet sequence homology to the ASGP receptors is confined exclusively to the carbohydrate-binding domain (specific for *N*-acetylglucosamine; Drickamer, 1987). This is also the case for the lymphocyte immunoglobulin E (IgE, $Fc_{\epsilon}RII$) receptor where exons 9–11 of its gene encode a putative carbohydrate-binding domain with two introns at exactly the same positions as in the RHL-1 gene (Suter et al., 1987; Figure 1A). The rest of the protein has no sequence homology to the ASGP receptors and an entirely different exon organization. It appears thus unlikely that the ASGP receptors, the chicken hepatic lectin, and the IgE receptor have evolved from a single ancestral gene; more likely, they have been assembled by "exon shuffling" from different donor genes such that they now share only the carbohydrate-binding domain. In fact, a homologous region consisting of 13 identities within the amino-terminal 42 amino acids of the chicken hepatic lectin and the IgE receptor suggests that these two genes share two additional exons with a common origin (2 and 3 in the IgE receptor gene; Figure 3).

Rearrangement of functional units from different ancestral genes is obvious in a growing family of proteins which contain a carbohydrate-binding domain homologous to that of the ASGP receptors. This domain is set in vastly different sequence contexts; for instance, in combination with collagen-like sequences in soluble serum proteins, with glycosaminoglycan attachment regions, or with EGF-like repeats in extracellular matrix proteins [reviewed by Drickamer (1987, 1988), Suter et al. (1987), Petersen (1988), and Thiel and Reid (1989)]. Of these proteins, the genes of the rat mannose-binding protein A and of the human pulmonary surfactant apoprotein are known (Figure 1A; White et al., 1985; Drickamer & McCreary, 1987). Interestingly, the carbohydrate-binding domain in these genes is encoded by a single exon that most likely has evolved by fusion of three exons like those found still separate in the ASGP and IgE receptors. Other exons in these genes code for collagen-like domains.

Carbohydrate-Binding Domain

The carbohydrate-binding domain has been functionally identified by partial proteolysis in both RHL-1 and chicken hepatic lectin: a carboxy-terminal clostripain fragment consisting of residues 134–283 of RHL-1 and a subtilisin fragment of residues 71–207 of the chicken hepatic lectin are sufficient to bind galactose and *N*-acetylgalactosamine, respectively (Hsueh et al., 1986; Chiaccia & Drickamer, 1984). The cleavage sites are very close to the position of intron 6 in the RHL-1 gene, suggesting that the segment encoded by exons 7–9 forms a compact domain while the preceding sequence might form a flexible neck accessible to proteolytic cleavage. In the IgE receptor, the site of proteolytic cleavage, which occurs naturally releasing a soluble IgE-binding factor, is also located immediately preceding the homologous carboxy-ter-

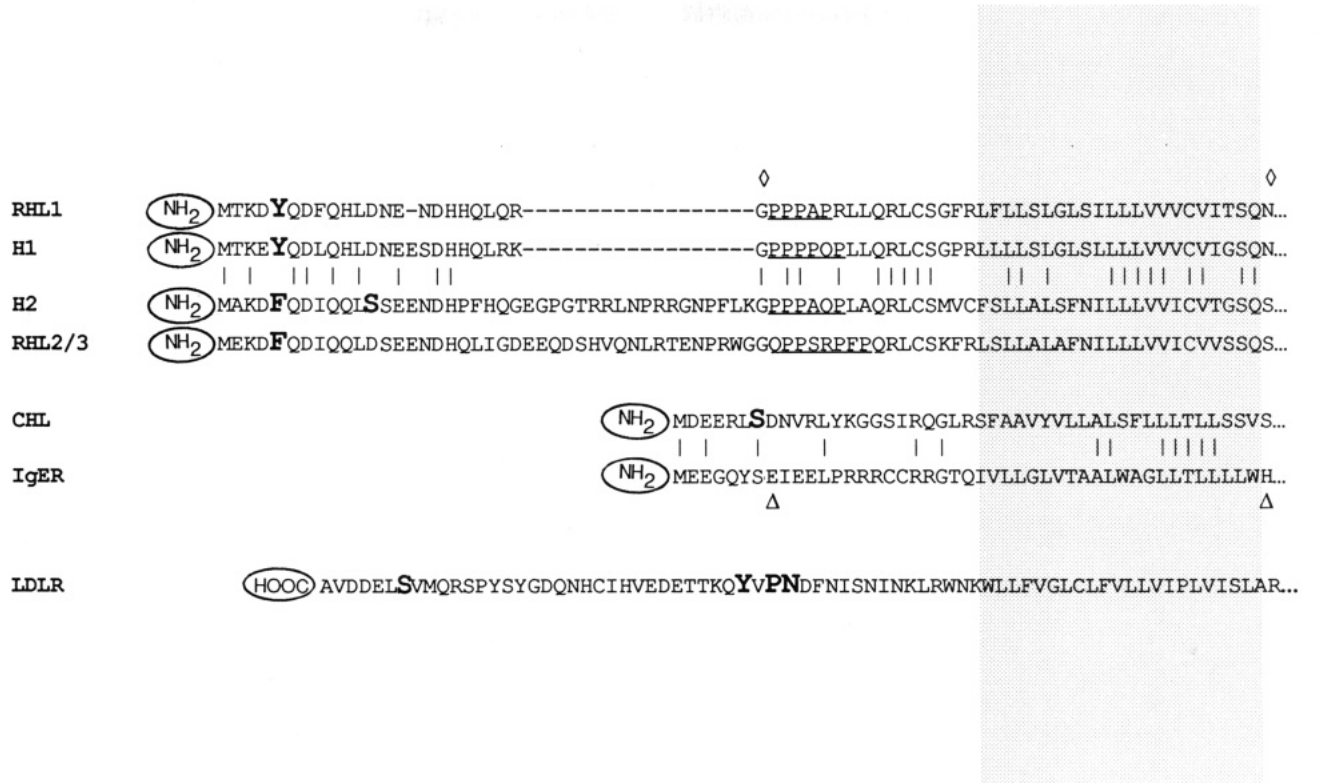

FIGURE 3: Sequences of the cytoplasmic and transmembrane domains of the rat and human ASGP receptors. Residues conserved in all four ASGP receptor sequences are indicated by lines. For comparison, the corresponding segments of the chicken hepatic lectin (CHL), the IgE receptor, and the LDL receptor are shown. Serine residues that are known to be phosphorylated as well as the residues implicated in endocytosis are in boldface. The proline-rich sequences in the ASGP receptor proteins are underlined, and the known exon boundaries are indicated by diamonds and triangles. The apolar phase of the lipid bilayer is stippled.

minal domain (Lüdin et al., 1987).

As mentioned above, homologous carbohydrate-binding domains have been discovered in many different proteins. For several of these proteins, there is no direct evidence for the binding of carbohydrates, and physiological ligands have not been identified. Carbohydrate-binding activity is inferred from the homology with the ASGP receptors. Surprisingly, even though the IgE receptor binds its ligand by the homologous domain (Bettler et al., 1989), it recognizes the protein and not the carbohydrate moiety of IgE (Vercelli et al., 1989). The hallmark of the known carbohydrate-binding proteins of this family is an absolute requirement for Ca^{2+}. To emphasize this feature, they have been classified as C-type animal lectins, as opposed to the S-type lectins that are characterized by essential free thiol groups (Drickamer, 1988) and to other lectins that belong to neither group (e.g., the mannose 6-phosphate receptors; Thiel & Reid, 1989). Comparison of the C-type lectin sequences reveals several invariant residues in a characteristic pattern (shown in Figure 1B). Among the invariant amino acids are four cysteines which were shown to form disulfide bridges in three proteins, the acorn barnacle lectin, echinoidin (a sea urchin lectin), and human tetranectin (Muramoto & Kamiya, 1986; Giga et al., 1987; Fuhlendorff et al., 1987). In a subclass of the C-type family, including the ASGP receptors, the IgE receptor, but not the chicken hepatic lectin, an additional two cysteines near the amino-terminal end of the domain are conserved. In tetranectin, echinoidin, and the barnacle lectin, these two cysteines have also been shown to be disulfide bonded.

The presence of intramolecular disulfide bonds in ASGP receptors was shown by the early observation that the rabbit receptor is readily inactivated by reducing agents (Kawasaki & Ashwell, 1977). In fact, differential sensitivity to β-mercaptoethanol has been employed to purify RHL-1 from the minor forms RHL-2/3, which lose binding activity under reducing conditions milder than those of the major form (Halberg et al., 1987). Furthermore, in vitro translated rat receptors acquire their functional conformation only when translocated into microsomes and under sufficiently oxidizing conditions (produced by addition of oxidized glutathione; Hsueh et al., 1986; Halberg et al., 1987), suggesting the formation of disulfide bridges involving microsomal disulfide isomerase.

An important feature of many endocytic receptors is the pH dependence of ligand binding which is essential for the efficient release of the ligand in acidic endosomes (Mellman et al., 1986). The ASGP receptors and the chicken hepatic lectin gradually lose ligand-binding activity at a pH lower than 6.5. For the rabbit ASGP receptors it was shown that lowering the pH from 7.4 to 5.6 is accompanied by changes in tryptophan fluorescence, efficiency of iodination, and sensitivity to papain digestion, suggesting a change in conformation (DiPaola & Maxfield, 1984). Similarly, the chicken hepatic lectin becomes more sensitive to proteolysis and displays an altered circular dichroism spectrum (Loeb & Drickamer, 1988). Loeb and Drickamer (1988) also showed for the chicken hepatic lectin that lowering the pH dramatically affects the affinity for Ca^{2+}: 2 mM Ca^{2+} is sufficient for half-maximal ligand binding at neutral pH, but approximately 30 mM Ca^{2+} is required at pH 5.4. Therefore, as the receptor–ligand complex enters the endosomal compartment, the acidic pH causes the release of both Ca^{2+} and ligand. Ca^{2+} can be replaced by Sr^{2+} and Ba^{2+} (in the order of efficiency $Sr^{2+} > Ca^{2+} > Ba^{2+}$) but not by Mg^{2+}. Ca^{2+}-binding studies performed with the chicken hepatic lectin and the rabbit ASGP receptor indicate that two molecules of Ca^{2+} are bound per polypeptide (Loeb &

Drickamer, 1988; Andersen et al., 1982).

Transmembrane Signal–Anchor Domain

For most secretory and transmembrane proteins studied, insertion into the endoplasmic reticulum (ER) membrane is initiated by an amino-terminal signal peptide of 16–26 residues with a polar, basic amino terminus and a central, hydrophobic domain of typically 7–12 apolar amino acids (Walter et al., 1984; Wickner & Lodish, 1985; Gierasch, 1989). This sequence is recognized by the signal recognition particle (SRP) and targeted to the ER by association with the SRP receptor in the microsomal membrane. As the nascent polypeptide is extruded through the bilayer, the signal, in most cases, is proteolytically removed by a specific signal peptidase in the ER lumen. Thus, a new amino terminus is generated on the exoplasmic side of the membrane. Translocation of membrane proteins is terminated by a hydrophobic stop–transfer sequence, consisting of approximately 20 apolar residues, which also serves as a helical transmembrane anchor. By this mechanism, single-spanning proteins (classified as type I membrane proteins) acquire an orientation with the amino terminus on the exoplasmic side and the carboxy terminus on the cytoplasmic side of the membrane. Examples among the endocytic receptors include those for LDL, mannose 6-phosphate, EGF, insulin, and poly(Ig).

The orientation of the ASGP receptors is just the opposite: the carboxy-terminal portion with the carbohydrate-binding domain and the glycosylation sites is accessible on the extracellular side, and the amino terminus is exposed to the cytoplasm. The beginning of the open reading frame of the cDNAs is identical with the amino-terminal sequences of the mature proteins except for the removal of the initiator methionine, indicating that there is no substantial proteolytic processing. The amino-terminal sequence is furthermore entirely hydrophilic and does not resemble any known cleaved or uncleaved signal sequences. The ASGP receptors were therefore predicted to contain an internal signal for membrane insertion. The membrane-spanning domain consisting of at least 20 uncharged, largely hydrophobic residues appeared to be the most likely candidate.

By use of an in vitro assay for membrane insertion, it could be shown that this transmembrane domain is both necessary and sufficient to target the protein to the ER membrane and to initiate translocation (Spiess & Lodish, 1986; Holland & Drickamer, 1986). The cDNAs of H1 and RHL-1 and of deletion and fusion constructs were transcribed in vitro, and the resulting mRNAs were translated in cell-free systems in the presence of dog pancreas microsomes as acceptor membranes. Insertion was judged by the glycosylation of the proteins and their resistance to exogenous protease. Membrane insertion was found to be abolished when the transmembrane domain was deleted; this domain alone, when fused to normally cytoplasmic proteins, such as tubulin or globin, initiated their translocation across the membrane. Insertion mediated by this internal signal occurs only cotranslationally and is dependent on SRP and the SRP receptor. Thus, the transmembrane domain behaves similarly to classical signal peptides but differs in three aspects: (i) It is located internally and initiates insertion only after the synthesis of a hydrophilic amino-terminal domain is already completed in the cytoplasm. This requires that, as the amino-terminal domain folds in the cytoplasm, it does not sterically interfere with signal recognition by SRP. A proline-rich region immediately prior to the transmembrane domain of the ASGP receptors conceivably serves the purpose of separating the preceding domain from the signal. The proline-rich sequence is encoded in the same exon as the hydrophobic sequence and thus might be part of the same functional unit. (ii) The internal signal lacks a signal peptidase cleavage site. However, a cryptic cleavage site at the very end of the hydrophobic segment (after Gly60) could be detected in a deletion construct of H1 lacking almost the entire amino-terminal domain (Schmid & Spiess, 1988). Most likely, the structure of the cytoplasmic domain affects the position of the receptor in the membrane and thus the accessibility of the potential cleavage site to signal peptidase. (iii) The apolar domain anchors the final protein by spanning the membrane. Like stop–transfer sequences, but unlike cleaved signals, it consists of 20 uncharged amino acids, corresponding to the length required to span the hydrophobic core of a biological membrane as an α-helix. By deletion analysis it was shown that nearly half the hydrophobic segment can be deleted from either end without abolishing insertion activity (Spiess & Handschin, 1987). An apolar stretch of 10–12 residues is sufficient for targeting, for insertion, and for generating a transmembrane orientation. However, the stability of membrane anchoring, as judged by extractability of the protein from the lipid membrane under alkaline conditions (pH 11.5), is very sensitive to the length of the apolar domain and is dramatically reduced even when only two residues are deleted.

A growing number of single-spanning membrane proteins share the topology (type II membrane proteins) and mode of insertion of the ASGP receptors, among them the chicken hepatic lectin, the receptors for IgE and transferrin (Zerial et al., 1986), and the macrophage scavenger receptors (Kodama et al., 1990). However, other proteins with internal signal sequences are inserted in the opposite orientation, i.e., their amino-terminal sequence is translocated across the membrane (e.g., cytochrome P-450 and opsin; Sakaguchi et al., 1987; Ovchinnikov, 1982). The final topology is not determined by the apolar segment of the signal but rather by the hydrophilic sequences flanking it. The cytoplasmic flanking sequences of transmembrane segments, including those of the ASGP receptors, are characteristically enriched in positively charged amino acids (the "positive inside rule"; von Heijne, 1986; von Heijne & Gavel, 1988). In addition, a striking correlation was found between the transmembrane orientation of signal–anchor sequences and the charge difference between the 15 carboxy-terminal and the 15 amino-terminal flanking residues (Hartmann et al., 1989). When the flanking charges of H1 were changed by site-directed mutagenesis, the majority of the protein expressed in transfected COS-7 cells indeed inserted in the opposite orientation (amino terminus exoplasmic, carboxy terminus cytoplasmic), indicating that the flanking charges of a signal sequence largely determine whether the amino- or the carboxy-terminal sequence is translocated across the membrane (Beltzer et al., submitted for publication).

Cytoplasmic Domain

The initial event in endocytosis is the clustering of receptors in coated pits. The receptor domains exposed to the cytoplasm appear to play a crucial role in the segregation of endocytic proteins. Deletion of the cytoplasmic domains of the receptors for LDL, transferrin, and poly(Ig) results in a dramatic reduction of the internalization rate (Lehrman et al., 1985; Rothenberger et al., 1987; Mostov et al., 1986). Furthermore, it was found that the number of coated pits observed in transfected cells correlated with the number of wild-type transferrin receptors expressed but not with the number of tail-less transferrin receptors, suggesting that the cytoplasmic tails catalyze the formation of coated pits (Iacopetta et al., 1988).

The principal component of the coat is clathrin [reviewed by Pearse (1987), Pearse and Crowther (1987), and Brodsky (1988)]: three heavy chains of 180 kDa and three light chains of 35 kDa form triskelions, which in turn associate in a lattice to form the outer shell of coated pits and vesicles. Associated on the inside are accessory or adaptor proteins. These are multiple polypeptides in the range of 100–115, 47–50, and 16–20 kDa, which assemble to complexes in a ratio of 2:1:1 (Manfredi & Bazari, 1987). Two types of adaptor complexes have been purified by hydroxylapatite chromatography, HA-I and HA-II, which are found in coated vesicles budding from the Golgi or from the plasma membrane, respectively. There is increasing evidence that these adaptor complexes recognize the cytoplasmic domain of receptors and induce the formation of coated pits. Pearse (1988) demonstrated a direct interaction of the cytoplasmic tail of the LDL receptor with HA-II but not HA-I complexes in vitro. Recently, we have been able to demonstrate an interaction of isolated human ASGP receptor with a 100-kDa adaptor protein of HA-II complexes after separation by gel electrophoresis and transfer to nitrocellulose (J. P. Beltzer, manuscript in preparation). The interaction is specific for the cytoplasmic domain since receptor binding can be competed with an excess of a bacterial fusion protein containing the amino-terminal domain of H1.

The cytoplasmic domains of endocytic receptors are very diverse in length and in sequence (Figure 3). In addition, some are amino terminal (e.g., those of the ASGP, transferrin, and IgE receptors) and others carboxy terminal [e.g., those of the LDL, poly(Ig), and mannose 6-phosphate receptors]. There are no obvious conserved recognition signals.

One of the natural LDL receptor mutations causing familial hypercholesterolemia is the result of a single amino acid change: tyrosine-807 was mutated to cysteine (Davis et al., 1986). Internalization is strongly reduced with any amino acid other than tyrosine, phenylalanine, and, to a lesser extent, tryptophan at this position (Davis et al., 1987). In addition, insertion of a tyrosine into the cytoplasmic domain of influenza hemagglutinin (in one of three positions tested) caused this normally resident surface protein to be endocytosed via coated vesicles (Lazarovits & Roth, 1988). Most endocytic receptors contain at least one tyrosine in their cytoplasmic tail, and its importance for endocytosis has been shown for the mannose 6-phosphate and transferrin receptors (Lobel et al., 1989; Jing et al., 1990). Of the ASGP receptors the major forms, H1 and RHL-1, contain a single tyrosine in position 5 (Figure 3). The minor forms lack tyrosines entirely but contain a phenylalanine at position 5. Mutation of tyrosine-5 to alanine in H1 resulted in a 3-fold reduction of the rate of internalization in transfected fibroblasts, suggesting that it is part of the signal for endocytosis (C. Fuhrer, manuscript in preparation). Whether phenylalanine-5 in H2 performs a similar function remains to be analyzed. Further mutagenesis of other residues surrounding the tyrosine of the LDL receptor tail led to the identification of an amino acid sequence, Asn-Pro-X-Tyr, that appears to be required for efficient uptake of LDL. It also occurs in the cytoplasmic domains of several other receptors (Chen et al., 1990) but has no similarity to the sequences of the ASGP receptor tails. The diversity of adaptor proteins might account for different types of recognition signals.

Phosphorylation of the cytoplasmic domains of membrane receptors is a general phenomenon that in several systems is known to affect receptor function [reviewed by Sibley et al. (1984)]. Recently, phosphorylation of the cytoplasmic tyrosine of the human ASGP receptor has been demonstrated in the presence of phosphatase inhibitors (Fallon, 1990). A functional significance for this modification still remains to be established. In addition, the rat and human ASGP receptors (like many other receptors, e.g., chicken hepatic lectin, transferrin, and IgE receptors) are phosphorylated on serine residues (Takahashi et al., 1985; Schwartz, 1984a; Drickamer & Mamon, 1982). Exogenous activation of protein kinase C by the addition of phorbol esters leads to hyperphosphorylation and, depending on the cell type, to a redistribution of the receptor. Surface ASGP receptors are hyperphosphorylated very rapidly after addition of phorbol esters to HepG2 cells. The number of surface receptors is reduced by 40–50% with a half-time of 20 min, and the majority of phosphorylated receptors accumulate intracellularly (Fallon & Schwartz, 1986, 1987, 1988). The rate of internalization and the total number of binding sites are not affected, but the phosphorylated receptor population appears not be recycled. A similar situation is found for the transferrin receptor. However, when its site of phosphorylation, serine-24, was mutated, receptor function and phorbol ester induced redistribution were unchanged, suggesting that serine phosphorylation is not essential for endocytosis and recycling and that there is no causal relationship between receptor phosphorylation and redistribution (Davis & Meisner, 1987; Rothenberger et al., 1987; McGraw et al., 1988). It is more likely that other substrates of protein kinase C affect cellular traffic in a more general way (Buys et al., 1984).

In agreement with these studies, the distribution of ASGP receptors H1 and H2 in transfected mouse fibroblast cells is not affected upon treatment with phorbol esters, even though they are hyperphosphorylated. In addition, a mutant of H1 lacking both cytoplasmic serine residues (at positions 16 and 37) is endocytosed and recycled normally in transfected fibroblasts (Geffen et al., unpublished results). It has been reported that the minor forms of the rat ASGP receptors, RHL-2 and RHL-3, are predominantly phosphorylated (Takahashi et al., 1985). Among the human receptors, H2 is also at least 10-fold more phosphorylated than H1. We have identified the phosphorylation site to be serine-12 (out of three serines at positions 12, 13, and 55; Geffen et al., unpublished results). Unlike typical sites for modification by protein kinase C, serine-12 is not preceded by positively charges residues. Since it is not conserved in RHL-2/3, other serines (at positions 13, 28, 46, and 55) must be phosphorylated in these proteins. The ubiquity and apparent specificity of receptor phosphorylation suggest a physiological function, which, however, for transport receptors is still entirely obscure.

Connecting Domain(s)

No function has been directly assigned to the segment connecting the cytoplasmic and transmembrane domains with the carbohydrate-binding domain in the ASGP receptors. This segment consists of 80 residues encoded by three exons (in the chicken hepatic lectin it is less than 40 residues). This portion might serve as a spacer or "neck" to expose the carbohydrate-binding domain on the cell surface. It could also be responsible for interactions with other receptor polypeptides to generate functional receptor oligomers. [For the chicken hepatic lectin the segment necessary for oligomerization has been narrowed to a segment between residue 11 and the carbohydrate-binding domain (Loeb & Drickamer, 1987).]

Perspectives

The availability of receptor cDNAs and cell culture expression systems now allows the functional analysis of subunits individually and in combination. Mutant cDNAs prepared in vitro by site-directed mutagenesis can be expressed in the

cell, and the behavior of the mutant and wild-type proteins can be analyzed. The aim is to identify and characterize the features or signals responsible for directing the receptor through the cellular compartments. Some of the experiments performed in this direction have already been discussed above. Expression of individual subunits revealed that association of the two ASGP receptor subunits is necessary for the efficient transport of both subunits to the plasma membrane and for high-affinity ligand binding. By use of a ligand-independent approach to monitor receptor internalization, it has been shown that subunit H1 alone is constutitively internalized and recyled with kinetics similar to those of the complete H1–H2 complex (Geffen et al., 1989). This was also shown for RHL-1 (Braiterman et al., 1989). The major subunit of the ASGP receptor thus contains all the signals necessary for endocytosis and recycling. This finding allows the mutational analysis of a single receptor subunit without the risk of a compensatory effect by the other subunit or the need to mutate the second subunit in the homologous position. With this approach, it could be demonstrated that tyrosine-5 of subunit H1 plays an important part in endocytosis and that serines at positions 16 and 37 do not (see above). Further studies will be required to obtain a more complete description of the sequence pattern necessary for internalization. Virtually nothing is known about the requirements for a protein to be recycled back to the plasma membrane or transported to lysosomes.

Additional sorting processes are operating in polarized cells when the protein is specifically transported to the basolateral plasma membrane domain in biosynthesis and recycling. To analyze this function, ASGP receptors have been expressed in Madin–Darby canine kidney (MDCK) cells, an epithelial cell line that, in contrast to hepatocytes, can be grown in culture as a polarized monolayer (Wessels et al., 1989; Graeve et al., 1989a). Both subunit H1 alone and the combination of RHL-1 and RHL-2 were found to be specifically transported to the basolateral plasma membrane, suggesting either that the same basolateral targeting signal is operative in hepatocytes and MDCK cells or that basolateral transport occurs "by default", i.e., without the requirement for a sorting signal. The rat subunits were shown to internalize ASGPs specifically from the basal side of the monolayer. These cell lines (as well as the single-subunit chicken hepatic lectin expressed in MDCK cells; Greave et al., 1989b) provide the basis for the characterization of the mechanisms that control targeting and recycling of proteins to the basolateral membrane of epithelial cells.

An important question that remains to be addressed, which pertains more specifically to the ASGP receptor, is how the two subunits interact to create a high-affinity carbohydrate-binding entity. The advantage that this more complex structure provides, e.g., for regulation of receptor function or specificity, is currently not clear. During the last decade, analysis of the ASGP receptor system has contributed to the description and understanding of receptor-mediated endocytosis. Future studies will focus on the molecular details of the individual processes involved.

ACKNOWLEDGMENTS

I thank Howard Riezman and my co-workers for critically reading the manuscript. Our own research has been supported by grants from the Swiss National Science Foundation.

REFERENCES

Amara, J. F., Lederkremer, G., & Lodish, H. F. (1989) *J. Cell Biol. 109*, 3315–3324.

Andersen, T. T., Freytag, J. W., & Hill, R. L. (1982) *J. Biol. Chem. 257*, 8036–8041.

Ashwell, G., & Harford, J. (1982) *Annu. Rev. Biochem. 51*, 531–554.

Baenziger, J. U., & Maynard, Y. (1980) *J. Biol. Chem. 255*, 4607–4613.

Bettler, B., Maier, R., Ruegg, D., & Hofstetter, H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7118–7122.

Bischoff, J., & Lodish, H. F. (1987) *J. Biol. Chem. 262*, 11825–11832.

Bischoff, J., Libresco, S., Shia, M. A., & Lodish, H. F. (1988) *J. Cell. Biol. 106*, 1067–1074.

Braiterman, L. T., Chance, S. C., Porter, W. R., Lee, Y. C., Townsend, R. R., & Hubbard, A. L. (1989) *J. Biol. Chem. 264*, 1682–1688.

Breitfeld, P. P., Simmons, C. F., Strous, G. J. A. M., Geuze, H. J., & Schwartz, A. L. (1985) *Int. Rev. Cytol. 97*, 47–95.

Brodsky, F. M. (1988) *Science 242*, 1396–1402.

Buys, S. S., Keogh, E. A., & Kaplan, J. (1984) *Cell 38*, 569–576.

Chen, W. J., Goldstein, J. L., & Brown, M. S. (1990) *J. Biol. Chem. 265*, 3116–3123.

Chiacchia, K. B., & Drickamer, K. (1984) *J. Biol. Chem. 259*, 15440–15446.

Ciechanover, A., Schwartz, A. L., & Lodish, H. F. (1983) *Cell 32*, 267–275.

Daniels, C. K., Schmucker, D. L., & Jones, A. L. (1989) *Hepatology 9*, 229–234.

Davis, C. G., Lehrman, M. A., Russell, D. W., Anderson, R. G. W., Brown, M. S., & Goldstein, J. L. (1986) *Cell 45*, 15–24.

Davis, C. G., van Driel, I. R., Russell, D. W., Brown, M. S., & Goldstein, J. L. (1987) *J. Biol. Chem. 262*, 4075–4082.

Davis, R. J., & Meisner, H. (1987) *J. Biol. Chem. 262*, 16041–16047.

DiPaola, M., & Maxfield, F. R. (1984) *J. Biol. Chem. 259*, 9163–9171.

Drickamer, K. (1981) *J. Biol. Chem. 256*, 5827–5839.

Drickamer, K. (1987) *Kidney Int. 32*, S167–S180.

Drickamer, K. (1988) *J. Biol. Chem. 263*, 9557–9560.

Drickamer, K., & Mamon, J. F. (1982) *J. Biol. Chem. 257*, 15156–15161.

Drickamer, K., & McCreary, V. (1987) *J. Biol. Chem. 262*, 2582–2589.

Drickamer, K., Mamon, J. F., Binns, G., & Leung, J. O. (1984) *J. Biol. Chem. 259*, 770–778.

Fallon, R. J. (1990) *J. Biol. Chem. 265*, 3401–3406.

Fallon, R. J., & Schwartz, A. L. (1986) *J. Biol. Chem. 261*, 15081–15089.

Fallon, R. J., & Schwartz, A. L. (1987) *Mol. Pharmacol. 32*, 348–355.

Fallon, R. J., & Schwartz, A. L. (1988) *J. Biol. Chem. 263*, 13159–13166.

Fuhlendorff, J., Clemmensen, I., & Magnusson, S. (1987) *Biochemistry 26*, 6757–6764.

Geffen, I., Wessels, H. P., Roth, J., Shia, M. A., & Spiess, M. (1989) *EMBO J. 8*, 2855–2862.

Gething, M.-J., McCammon, K., & Sambrook, J. (1986) *Cell 46*, 939–950.

Geuze, H. J., Slot, J. W., Strous, G. J. A. M., Lodish, H. F., & Schwartz, A. L. (1982) *J. Cell Biol. 92*, 865–870.

Geuze, H. J., Slot, J. W., Strous, G. J., Lodish, H. F., & Schwartz, A. L. (1983a) *Cell 32*, 277–287.

Geuze, H. J., Slot, J. W., Strous, G. J., & Schwartz, A. L. (1983b) *Eur. J. Cell Biol. 32*, 38–44.

Geuze, H. J., Slot, J. W., Strous, G. J., Luzio, J. P., & Schwartz, A. L. (1984) *EMBO J. 3*, 2677–2685.
Gierasch, L. M. (1989) *Biochemistry 28*, 934–930.
Giga, Y., Ikai, A., & Takahashi, K. (1987) *J. Biol. Chem. 262*, 6197–6203.
Goldstein, J. L., Brown, M. S., Anderson, R. G. W., Russell, D. W., & Schneider, W. J. (1985) *Annu. Rev. Cell Biol. 1*, 1–39.
Graeve, L., Drickamer, K., & Rodriguez-Boulan, E. (1989) *J. Cell Biol. 109*, 2809–2816.
Graeve, L., Patzak, A., Drickamer, K., & Rodriguez-Boulan, E. (1990) *J. Biol. Chem. 265*, 1216–1224.
Halberg, D. F., Wager, R. E., Farrell, D. C., Hildreth, J., Quesenberry, M. S., Loeb, J. A., Holland, E. C., & Drickamer, K. (1987) *J. Biol. Chem. 262*, 9828–9838.
Hardy, M. R., Townsend, R. R., Parkhurst, S. M., & Lee, Y. C. (1985) *Biochemistry 24*, 22–28.
Hartmann, E., Rapoport, T. A., & Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5786–5790.
Herzig, M. C., & Weigel, P. H. (1989) *Biochemistry 28*, 600–610.
Holland, E. C., & Drickamer, K. (1986) *J. Biol. Chem. 261*, 1286–1292.
Holland, E. C., Leung, J. O., & Drickamer, K. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7338–7342.
Hong, W., Le, A. V., & Doyle, D. (1988) *Hepatology 8*, 553–558.
Hsueh, E. C., Holland, E. C., Carrera, G. M. J., & Drickamer, K. (1986) *J. Biol. Chem. 261*, 4940–4947.
Hubbard, A. L., & Stukenbrok, H. (1979) *J. Cell Biol. 83*, 56–81.
Hubbard, A. L., Wilson, G., Ashwell, G., & Stukenbrok, H. (1979) *J. Cell Biol. 83*, 47–55.
Iacopetta, B. J., Rothenberger, S., & Kühn, L. C. (1988) *Cell 54*, 485–489.
Jing, S. Q., Spencer, T., Miller, K., Hopkins, C., & Trowbridge, I. S. (1990) *J. Cell Biol. 110*, 283–294.
Kawasaki, T., & Ashwell, G. (1976) *J. Biol. Chem. 251*, 1296–1302.
Kawasaki, T., & Ashwell, G. (1977) *J. Biol. Chem. 252*, 6536–6543.
Kodama, T., Freeman, M., Rohrer, L., Zabrecky, J., Matsudaira, P., & Krieger, M. (1990) *Nature 343*, 531–535.
Kreis, T. E., & Lodish, H. F. (1986) *Cell 46*, 929–937.
Lazarovits, J., & Roth, M. (1988) *Cell 53*, 743–752.
Lee, R. T., & Lee, Y. C. (1986) *Biochemistry 25*, 6835–6841.
Lee, R. T., & Lee, Y. C. (1987) *Biochemistry 26*, 6320–6329.
Lee, Y. C., Townsend, R. R., Hardy, M. R., Lonngren, J., Arnarp, J., Haraldsson, M., & Lonn, H. (1983) *J. Biol. Chem. 258*, 199–202.
Lehrman, M. A., Goldstein, J. L., Brown, M. S., Russell, D. W., & Schneider, W. J. (1985) *Cell 41*, 735–743.
Leung, J. O., Holland, E. C., & Drickamer, K. (1985) *J. Biol. Chem. 260*, 12523–12527.
Lobel, P., Fujimoto, K., Ye, R. D., Griffiths, G., & Kornfeld, S. (1989) *Cell 57*, 787–796.
Loeb, J. A., & Drickamer, K. (1987) *J. Biol. Chem. 262*, 3022–3029.
Loeb, J. A., & Drickamer, K. (1988) *J. Biol. Chem. 263*, 9752–9760.
Lüdin, C., Hofstetter, H., Sarfati, M., Levy, C. A., Suter, U., Alaimo, D., Kilchherr, E., Frost, H., & Delespesse, G. (1987) *EMBO J. 6*, 109–114.
Lunney, J., & Ashwell, G. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 341–343.
Manfredi, J. J., & Bazari, W. L. (1987) *J. Biol. Chem. 262*, 12182–12188.
Matsuura, S., Nakada, H., Sawamura, T., & Tashiro, Y. (1982) *J. Cell Biol. 95*, 864–875.
May, W. S., & Cuatrecasas, P. (1985) *J. Membr. Biol. 88*, 205–215.
McGraw, T. E., Dunn, K. W., & Maxfield, F. R. (1988) *J. Cell Biol. 106*, 1061–1066.
McPhaul, M., & Berg, P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8863–8867.
McPhaul, M., & Berg, P. (1987) *Mol. Cell. Biol. 7*, 1841–1847.
Mellman, I., Fuchs, R., & Helenius, A. (1986) *Annu. Rev. Biochem. 55*, 663–700.
Mostov, K. E., & Simister, N. E. (1985) *Cell 43*, 389–390.
Mostov, K. E., de Bruyn Kops, A., & Deitcher, D. L. (1986) *Cell 47*, 359–364.
Mueller, S. C., & Hubbard, A. L. (1986) *J. Cell Biol. 102*, 932–942.
Muramoto, K., & Kamiya, H. (1986) *Biochim. Biophys. Acta 874*, 285–295.
Ovchinnikov, Y. A. (1982) *FEBS Lett. 148*, 179–191.
Patthy, L. (1988) *J. Mol. Biol. 202*, 689–696.
Pearse, B. M. F. (1987) *EMBO J. 6*, 2507–2512.
Pearse, B. M. F. (1988) *EMBO J. 7*, 3331–3336.
Pearse, B. M. F., & Crowther, R. A. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 49–68.
Petell, J. K., & Doyle, D. (1985) *Arch. Biochem. Biophys. 241*, 550–560.
Petersen, T. E. (1988) *FEBS Lett. 231*, 51–53.
Regoeczi, E., Hatton, M. W. C., & Charlwood, P. A. (1975) *Nature 254*, 699–701.
Rothenberger, S., Iacopetta, B. J., & Kühn, L. C. (1987) *Cell 49*, 423–431.
Sakaguchi, M., Mihara, K., & Sato, R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3361–3364.
Sanford, J. P., Elliott, R. W., & Doyle, D. (1988) *DNA 7*, 721–728.
Sawamura, T., Nakada, H., Hazama, H., Shiozaki, Y., Sameshima, Y., & Tashiro, Y. (1984) *Gastroenterology 87*, 1217–1221.
Sawyer, J. T., Sanford, J. P., & Doyle, D. (1988) *J. Biol. Chem. 263*, 10534–10538.
Schlessinger, J. (1988) *Biochemistry 27*, 3119–3123.
Schmid, S. R., & Spiess, M. (1988) *J. Biol. Chem. 263*, 16886–16891.
Schwartz, A. L. (1984a) *Biochem. J. 223*, 481–486.
Schwartz, A. L. (1984b) *CRC Crit. Rev. Biochem. 16*, 207–233.
Schwartz, A. L., & Rup, D. (1983) *J. Biol. Chem. 258*, 11249–11255.
Schwartz, A. L., Fridovich, S. E., & Lodish, H. F. (1982) *J. Biol. Chem. 257*, 4230–4237.
Schwartz, A. L., Steer, C. J., & Kempner, E. S. (1984) *J. Biol. Chem. 259*, 12025–12029.
Shia, M. A., & Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1158–1162.
Sibley, D. R., Benovic, J. L., Caron, M. G., & Lefkowitz, R. J. (1987) *Cell 48*, 913–922.
Soutar, A. K. (1989) *Nature 341*, 106–107.
Spiess, M., & Lodish, H. F. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6465–6469.
Spiess, M., & Lodish, H. F. (1986) *Cell 44*, 177–185.
Spiess, M., & Handschin, C. (1987) *EMBO J. 6*, 2683–2691.

Spiess, M., Schwartz, A. L., & Lodish, H. F. (1985) *J. Biol. Chem. 260*, 1979–1982.
Steer, C. J., Kempner, E. S., & Ashwell, G. (1981) *J. Biol. Chem. 256*, 5851–5856.
Stoorvogel, W., Geuze, H. J., Griffith, J. M., Schwartz, A. L., & Strous, G. J. (1989) *J. Cell Biol. 108*, 2137–2148.
Südhof, T. C., Goldstein, J. L., Brown, M. S., & Russell, D. W. (1985) *Science 228*, 815–822.
Suter, U., Bastos, R., & Hofstetter, H. (1987) *Nucleic Acids Res. 15*, 7295–7308.
Takahashi, T., Nakada, H., Okumura, T., Sawamura, T., & Tashiro, Y. (1985) *Biochem. Biophys. Res. Commun. 126*, 1054–1060.
Thiel, S., & Reid, K. B. M. (1989) *FEBS Lett. 250*, 78–84.
Tolleshaug, H., & Berg, T. (1979) *Biochem. Pharmacol. 28*, 2919–2922.
Tycko, B., Keith, C. H., & Maxfield, F. R. (1983) *J. Cell Biol. 255*, 5971–5978.
Van Lenten, L., & Ashwell, G. (1972) *J. Biol. Chem. 247*, 4633–4638.
Vercelli, D., Helm, B., Marsh, P., Padlan, E., Geha, R. S., & Gould, H. (1989) *Nature 338*, 649–650.
von Heijne, G. (1986) *EMBO J. 5*, 3021–3027.
von Heijne, G., & Gavel, Y. (1988) *Eur. J. Biochem. 174*, 671–678.
Wall, D. A., & Hubbard, A. (1981) *J. Cell Biol. 90*, 687–696.
Walter, P., Gilmore, R., & Blobel, G. (1984) *Cell 38*, 5–8.
Watts, C. (1985) *J. Cell Biol. 100*, 633–637.
Weigel, P. H. (1980) *J. Biol. Chem. 255*, 6111–6120.
Wessels, H. P., Geffen, I., & Spiess, M. (1989) *J. Biol. Chem. 264*, 17–20.
White, R. T., Damm, D., Miller, J., Spratt, K., Schilling, J., Hawgood, S., Benson, B., & Cordell, B. (1985) *Nature 317*, 361–363.
Wickner, W. T., & Lodish, H. F. (1985) *Science 230*, 400–407.
Yarden, Y., & Ullrich, A. (1988) *Biochemistry 27*, 3113–3119.
Zerial, M., Melancon, P., Schneider, C., & Garoff, H. (1986) *EMBO J. 5*, 1545–1550.

Accelerated Publications

X-ray Crystal Structure of the Protease Inhibitor Domain of Alzheimer's Amyloid β-Protein Precursor[†,‡]

Thomas R. Hynes,* Michael Randal, Laura A. Kennedy, Charles Eigenbrot, and Anthony A. Kossiakoff
Protein Engineering Department, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143
Received July 31, 1990; Revised Manuscript Received September 4, 1990

ABSTRACT: Alzheimer's amyloid β-protein precursor contains a Kunitz protease inhibitor domain (APPI) potentially involved in proteolytic events leading to cerebral amyloid deposition. To facilitate the identification of the physiological target of the inhibitor, the crystal structure of APPI has been determined and refined to 1.5-Å resolution. Sequences in the inhibitor–protease interface of the correct protease target will reflect the molecular details of the APPI structure. While the overall tertiary fold of APPI is very similar to that of the Kunitz inhibitor BPTI, a significant rearrangement occurs in the backbone conformation of one of the two protease binding loops. A number of Kunitz inhibitors have similar loop sequences, indicating the structural alteration is conserved and potentially an important determinant of inhibitor specificity. In a separate region of the protease binding loops, APPI side chains Met-17 and Phe-34 create an exposed hydrophobic surface in place of Arg-17 and Val-34 in BPTI. The restriction this change places on protease target sequences is seen when the structure of APPI is superimposed on BPTI complexed to serine proteases, where the hydrophobic surface of APPI faces a complementary group of nonpolar side chains on kallikrein A versus polar side chains on trypsin.

The major component of cerebral amyloid deposits in Alzheimer's disease is amyloid β-protein (AP), a proteolytically derived peptide of amyloid β-protein precursor (APP), an integral membrane glycoprotein (Glenner & Wong, 1984). Considerable work has focused on the nature of the proteolytic events leading to the accumulation of AP (Sisodia et al., 1990) following the discovery that alternative splicing of the APP message yields three products, two containing a Kunitz protease inhibitor domain (Kitaguchi et al., 1988; Ponte et al., 1988; Tanzi et al., 1988). Recent studies indicate that the inhibitor-containing species is overproduced in Alzheimer's patients (Johnson et al., 1989; Tanaka et al., 1989). It has been suggested that APPI may block the normal degradation of APP, leading to the deposition of cerebral amyloid (Müller-Hill & Beyreuther, 1989; Selkoe, 1990). The site of normal proteolytic cleavage occurs at a Lys residue within the AP peptide sequence (Esch et al., 1990). The primary specificity determining residue in APPI is an Arg, indicating that APPI blocks proteases that cleave at basic residues. The inhibition spectrum of APPI toward a range of serine proteases has been shown to be broad but distinct from basic pancreatic trypsin inhibitor (BPTI) (Kitaguchi et al., 1990).

[†] This work was supported by NIH Grant GM33571.
[‡] The coordinates of APPI have been deposited in the Brookhaven Protein Data Bank (entry name 1AAP).

0006-2960/90/0429-10018$02.50/0