Molecular Mechanism of Sulfated Carbohydrate Recognition:
Structural and Biochemical Studies of the Cysteine-rich Domain of Mannose Receptor

Thesis by

Yang Liu

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology
Pasadena, California

2001

(Defended on September 14, 2000)

Acknowledgments

First, I would like to thank my advisor, Pamela J. Bjorkman, for her invaluable guidance, encouragement, and support throughout my graduate study at Caltech. I have tremendously benefited from her broad knowledge and insightful advice. In addition, I would like to thank my thesis committee members, Raymond Deshaies, Stephen Mayo, Douglas Rees and Ellen Rothenberg, for their advice and support.

I'd like to thank many members of Bjorkman, Rees, and Mayo groups for their help. In particular, my thanks go to Arthur Chirino for detailed instruction for crystallography; to Peter Snow and Inderjit Nangiana for expression of Ly49 molecules; to Gary Hathaway and Jie Zhou for assistance of mass spectrophotometry; and to Melanie Bennett, Tara Chapman, Hsiu-Ju Chiu, Sheng Ding, Caroline Enns, Tony Giannetti, Hank Hogan, Zsuzsi Hamburger, Jim Harber, Astrid Heikema, Andrew Herr, Jennifer Johnson, Jose Lebrón, Lynda Llamas, Warham Lance Martin, Christine Milburn, Chantal Morgan, Marta Murphy, David Penny, Malini Raghavan, Tirunelveli S. Ramalingam, Luis Sánchez, Huazhang Shen, Julia Shifman, Elizabeth Sprague, Xiaodong Su, Sarah Sun, Daniel Vaughn, Anthony West, Clinton White, Benjamin Willcox, Michael Williamson, Andrew P. Yeh and many others for various help.

My thanks must go out to my collaborators, who provided generous help for my research. They are Ziva Misulovin, Christine Leteux, Thomas Hanke, Hisao Takizawa, Christopher McMahon, David H. Raulet, Ten Feizi, and Michel C. Nussenzweig.

I am grateful to my long time friends, Hong Yang, Wei Lin, Hai Wang, and Tianhe Sun, for their friendship; to my relatives, Aimin Zhang, Yi Liu and many others, for their support; and to my parents, Shuxiang Liu and Junrong Liu, for their love and support. At last, my special thanks go to my wife, Lingling, for her love, patience, and encouragement.

Abstract

Mannose receptor (MR) is widely expressed on macrophages, immature dendritic cells, and a variety of epithelial and endothelial cells. It is a 180 kD type I transmembrane receptor whose extracellular region consists of three parts: the amino-terminal cysteine-rich domain (Cys-MR); a fibronectin type II-like domain; and a series of eight tandem C-type lectin carbohydrate recognition domains (CRDs). Two portions of MR have distinct carbohydrate recognition properties: Cys-MR recognizes sulfated carbohydrates and the tandem CRD region binds terminal mannose, fucose, and N-acetyl-glucosamine (GlcNAc). The dual carbohydrate binding specificity allows MR to interact with sulfated and non-sulfated polysaccharide chains, and thereby facilitating the involvement of MR in immunological and physiological processes. The immunological functions of MR include antigen capturing (through binding non-sulfated carbohydrates) and antigen targeting (through binding sulfated carbohydrates), and the physiological roles include rapid clearance of circulatory luteinizing hormone (LH), which bears polysaccharide chains terminating with sulfated and non-sulfated carbohydrates.

We have crystallized and determined the X-ray structures of unliganded Cys-MR (2.0 Å) and Cys-MR complexed with different ligands, including Hepes (1.7 Å), $4SO_4$-N-Acetylgalactosamine ($4SO_4$-GalNAc; 2.2 Å), 3SO4-Lewis^x (2.2 Å), 3SO4-Lewis^a (1.9 Å), and $6SO_4$-GalNAc (2.5 Å). The overall structure of Cys-MR consists of 12 anti-parallel β-strands arranged in three lobes with approximate three fold internal symmetry. The structure contains three disulfide bonds, formed by the six cysteines in the Cys-MR sequence. The

ligand-binding site is located in a neutral pocket within the third lobe, in which the sulfate group of ligand is buried. Our results show that optimal binding is achieved by a carbohydrate ligand with a sulfate group that anchors the ligand by forming numerous hydrogen bonds and a sugar ring that makes ring-stacking interactions with Trp117 of Cys-MR. Using a fluorescence-based assay, we characterized the binding affinities between Cys-MR and its ligands, and rationalized the derived affinities based upon the crystal structures. These studies reveal the mechanism of sulfated carbohydrate recognition by Cys-MR and facilitate our understanding of the role of Cys-MR in MR recognition of its ligands.

Table of Contents:

Chapter 1:

Introduction

For many interactions at the cell surface, selective recognition of specific carbohydrate structures is mediated by cognate receptors, designated as lectins. Functions of lectins are widely variable (Drickamer and Taylor, 1993). Some play important roles in the host innate immune system; some function to sort newly synthesized glycoproteins in cytoplasm; and some mediate endocytosis of selective subsets of circulating glycoproteins.

Mannose receptor (MR) contains more than one type of lectin domain (Stahl and Ezekowitz, 1998). First, MR has multiple C-type lectin-like domains, which allow binding of high mannose-containing carbohydrates with high avidity. Second, MR also contains another lectin domain, called cysteine-rich domain, which binds sulfated carbohydrates. The two types of lectin domain collaborate to bind ligands, thus providing MR the flexibility to interact with a diverse group of ligands. MR is widely distributed on a variety of cells throughout the body and has diverse functions. Here I will review the properties of MR and indicate the key roles this receptor plays in immune system as well as in physiological processes.

Properties of MR

Cellular distribution

MR was first identified in rat liver Kupffer cells (Schlesinger et al., 1978). Later, studies demonstrated that MR was also expressed on macrophages (Ezekowitz et al., 1990; Taylor et al., 1990), immature dendritic cells (Sallusto et al., 1995), hepatic

endothelial cells (Bijsterbosch et al., 1996; Magnusson and Berg, 1993), kidney mesangial cells (Liu et al., 1996), and retinal pigment epithelial cells (Shepherd et al., 1991). Expression of MR appears to be modulated on hematopoietic cells. On macrophages, MR expression is restricted to mature populations, and is regulated by cytokines (Mokoena and Gordon, 1985; Stein et al., 1992). On dendritic cells, MR is only expressed on immature dendritic cells (Caux et al., 1997; Sallusto et al., 1995). On the non-hematopoietic cells, such as hepatic endothelial cells and Kupffer cells, MR is constitutively expressed (Fiete et al., 1991).

Structural and functional characteristics

MR is a 180 KD transmembrane protein (Ezekowitz et al., 1990; Taylor et al., 1990) that contains three extracellular regions: 1) a membrane-distal amino-terminal cysteine-rich domain (Cys-MR); 2) a fibronectin type II-like domain (FnII); and 3) eight tandem C-type lectin carbohydrate-recognition domains (CRDs, (Stahl and Ezekowitz, 1998)) (Figure 1A). The extracellular region is linked to a small cytoplasmic domain via a transmembrane domain. MR has two distinct carbohydrate binding properties that are conferred by two different regions. The N-terminal Cys-MR domain binds sulfated carbohydrates (Fiete et al., 1998), while the tandem CRD region can bind non-sulfated carbohydrates (Taylor et al., 1992). The dual carbohydrate binding specificity provides MR the capability to bind ligands by interacting with their sulfated and non-sulfated oligosaccharide chains.

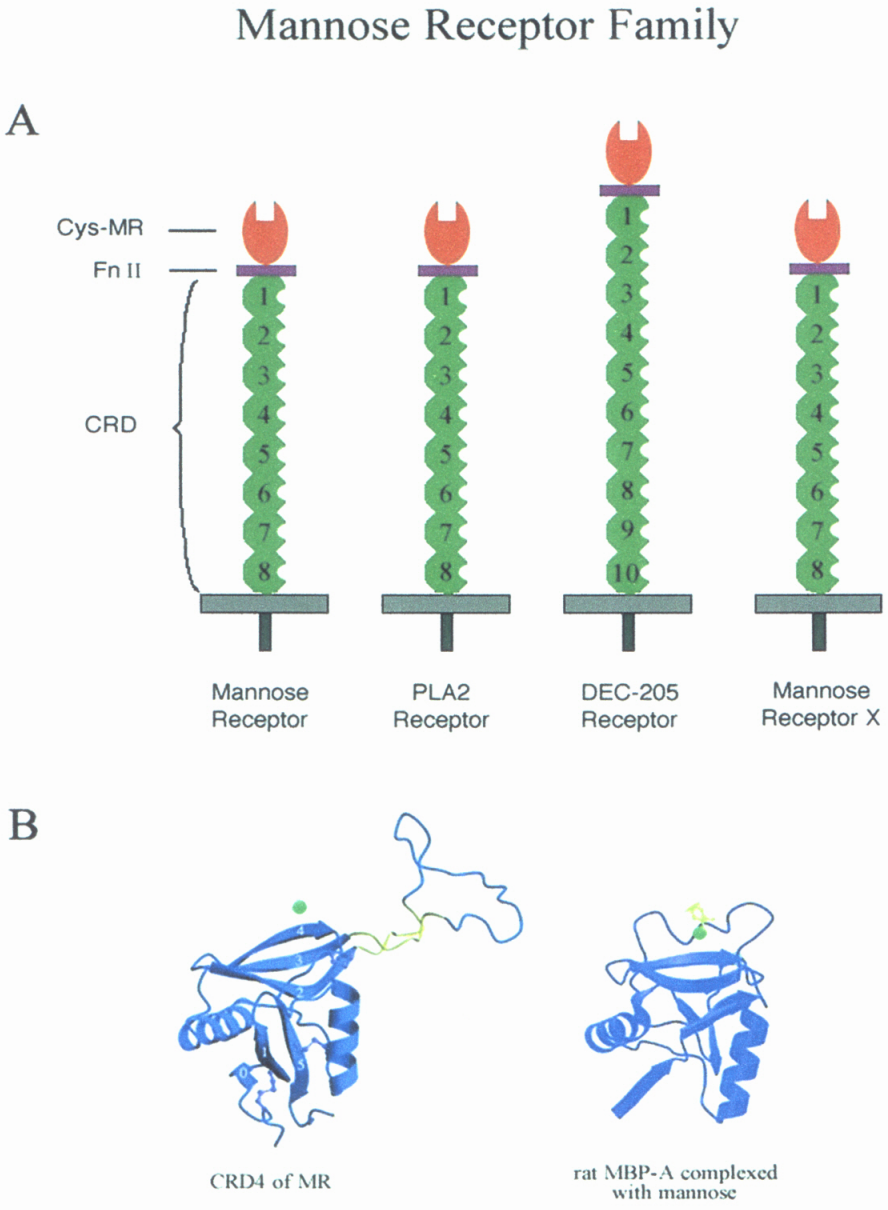


Figure 1. The mannose receptor (MR) family

(A) Schematic representation of the structures of MR and other members of the MR family shows common structural characteristics. CRDs are numbered from the membrane distal CRD. (Modified from Stahl et al., 1998.)
(B) Structures of the CRD4 from MR and the rat MBP-A complexed with mannose. Ca^{2+} is shown as a green sphere, and mannose is shown as yellow ball-and-stick. (Modified from Feinberg et al., 2000.)

Cys-MR has a total of six cysteines in a domain consisting of about 140 amino acids. Biochemical studies have shown that Cys-MR alone is able to bind sulfated carbohydrates (Fiete et al., 1998; Martinez-Pomares et al., 1999). Ligands of Cys-MR include N-glycans terminating with $4SO_4$-GalNAc on pituitary hormones (Fiete et al., 1998), chondroitin 4-sulfate chains terminating with $4SO_4$-GalNAc on hematopoietic cells, and sulfated blood group chains terminating with $3SO_4$-Gal, such as $3SO_4$-Lewis^x and $3SO_4$-Lewis^a (Leteux et al., 2000). To date, Cys-MR is one of the few known lectins that can specifically recognize sulfated carbohydrates. This special carbohydrate binding property designates Cys-MR as a novel lectin.

Several of the eight tandem C-type lectin CRDs of MR are involved in calcium-dependent binding to terminal mannose, fucose, or N-acetylglucosamine (GlcNAc) residues on the surface of pathogens or on the oligosaccharides of endogenous glycoproteins (Taylor et al., 1992; Taylor and Drickamer, 1993). Of the eight CRDs, only CRD4 has shown to bind mannose, fucose, or GlcNAc when expressed in isolation (Taylor et al., 1992). The structure of CRD4 has been determined (Feinberg et al., 2000), and it shares the same fold as the structure of rat serum mannose-binding protein (MBP-A), known as the representative C-type lectin structure (Weis et al., 1992) (Figure 1B). It is known that MR can bind mannosylated ligands with high binding affinity (K_D in the nM range) (Otter et al., 1992). However, with a K_D for monosaccharides in millimolar range (Mullin et al., 1997), CRD4 alone is not enough to mediate the high-affinity binding of MR. Other CRDs must contribute to the oligosaccharide binding. Studies have

shown that CRDs 4-8 must be present together to generate the binding affinity of the intact MR for natural ligands (Taylor and Drickamer, 1993). The requirement of multiple CRDs for high-affinity binding is reminiscent of multi-valent binding, a common feature of C-type and other lectins (Weis and Drickamer, 1996). The multi-valent binding theory is also supported by studies on oligosaccharide ligands of MR. It has been shown that MR preferentially binds oligo-mannoses that have branched rather than linear structures, such as core mannose on glycoproteins and yeast mannan (Kery et al., 1992) (Shibata et al., 1989). Overall, the multi-valent binding of CRDs 4-8 provides MR a special capability to recognize host-derived high mannose oligosaccharides as well as pathogenic high mannose structures.

MR is known to be internalized with high efficiency (Stahl et al., 1980). The cytoplasmic domain of MR contains an endocytic signal motif that is critical for MR mediated endocytosis and phagocytosis (Wu et al., 1996). By site-direct mutagensis of the cytoplasmic region of an MR-like PLA2 receptor, a sequence motif (NxxY) has been identified that encodes the major endocytic signal (Nicolas et al., 1995). This motif is conserved among the MR family members (Wu et al., 1996). Its existence on MR is thought to play a key role for MR-mediated endocytosis and phagocytosis.

MR family

MR is the prototypic member of a family of multilectin receptors named the MR family. Members of the family share similar structural characteristics. Other members of the MR family include the phospholipase A2 receptor (Ancian et al., 1995; Ishizaki et al.,

1994; Lambeau et al., 1994), the dendritic cell receptor DEC-205 (Jiang et al., 1995), and an MR-like receptor (Wu et al., 1996) (Figure 1A). Among the family members, MR is the only one that has two distinct carbohydrate recognition properties. Other members, though sharing sequence similarity (from 25% to 34% sequence identity), have not been proven to bind carbohydrates, although some have been predicted to bind carbohydrate (Wu et al., 1996). Here, I will review the current understanding of MR and its functions in both physiological and immunological processes.

Physiological roles of MR

Carbohydrate chains on glycoproteins are known to play a key role for the survival of glycoprotein. Clearance of glycoproteins *in vivo* has been shown to be a sugar-specific and receptor-mediated process (Ashwell and Morell, 1974). A well-understood pathway for clearance of galactose terminating glycoproteins (asialoglycoproteins) is mediated by asialoglycoprotein receptor on liver cells (Ashwell and Morell, 1974; Stryer, 1995). MR, abundantly expressed on macrophages and on hepatic endothelial cells, plays two important roles in maintaining homeostasis of glycoproteins. First, it participates in rapid removal of glycoproteins with terminal mannose and/or GlcNAc on their carbohydrate chains (Stahl et al., 1978). Second, it plays a critical role in clearance of pituitary hormones that bear sulfated oligosaccharide chains (Fiete et al., 1997; Fiete et al., 1991).

The role of MR in glycoprotein homeostasis

MR plays an important role in clearance of host-derived glycoproteins. Its homeostatic function includes uptake of glycoproteins bearing high mannose oligosaccharides, including tissue plasminogen activator (Otter et al., 1992; Smedsrod et al., 1988), lysosomal hydrolases (Stahl et al., 1978) and peroxidases (Shepherd and Hoidal, 1990). On liver endothelial cells, MR is involved in clearance of circulatory glycoproteins, such as tissue plasminogen activator. It was shown that endocytosis of tissue plasminogen activator is mediated by high affinity binding (apparent K_D = 3.5 nM) via MR in a mannose-specific manner (Otter et al., 1992; Smedsrod et al., 1988). Macrophage MR, on the other hand, participates in clearance of locally released glycoproteins. Both lysosomal hydrolases and peroxidases, released by degranulating neutrophils and macrophages in the process of inflammation, are potentially harmful enzymes to the host (Goldstein, 1977; Shepherd and Hoidal, 1990). Rapid clearance of them is mediated by MR on macrophages through MR binding to high mannose oligosaccharides on these enzymes (Stahl et al., 1978) (Shepherd and Hoidal, 1990; Shepherd et al., 1985). In this aspect, MR might play a role in the resolution phase of inflammation.

The role of MR in hormone regulation

Hormones are key coordinators in the endocrine systems in mammals. Synthesized by specific glands and tissues, hormones are transported to their action sites

via the bloodstream. They affect their target organs or tissues through binding to specific receptors on the target and thereby triggering a cascade of intracellular signals. Physiological functions under hormonal control include regulation of metabolic rates (growth hormone), control of sex activity and reproductive cycles (sex hormones), control of blood sugar (insulin), and modification of behavior (various hormones) (Randall et al., 1997; West, 1991).

For the regulation of sex activity and reproductive cycles, the pituitary gland is the primary organ for secretion of glycoprotein sex hormones, including luteinizing hormone (LH) and follicle-stimulating hormone (FSH). The secretion of LH and FSH is controlled by gonadotropin-releasing hormone, which is released by neurosecretory cells in the hypothalamus. Both LH and FSH are also called gonadotropins, since they act on gonads. They promote production and secretion of steroid sex hormones (progesterone and estrogens in females and testosterone in males) that are critical for development of primary and secondary sexual characteristics in both sexes (Randall et al., 1997; West, 1991) (Figure 2).

1) Structural characteristics and biological functions of gonadotropins

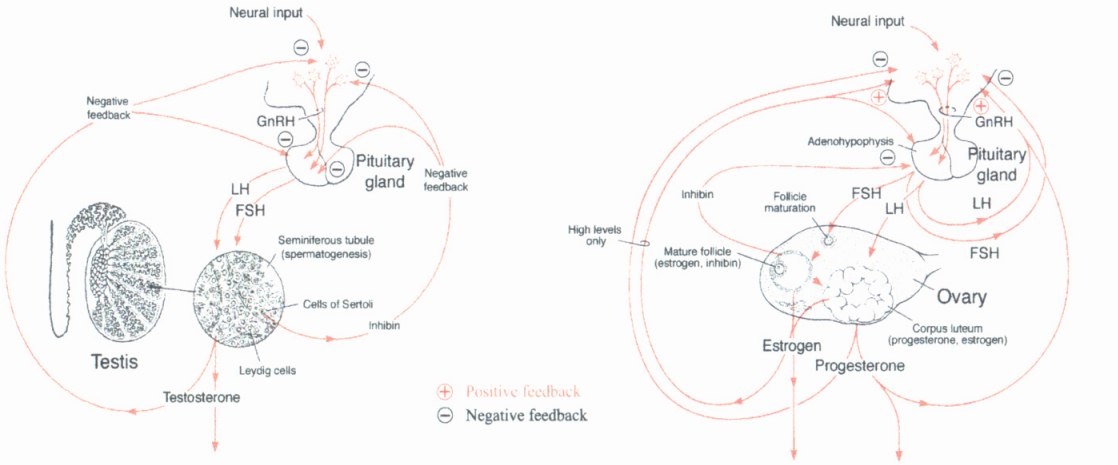
The gonadotropins (LH and FSH) belong to a family of glycoprotein hormones, which has two additional members, thyroid stimulating hormone (TSH) and choronic gonadotropin (CG). The glycoprotein hormones are heterodimers that consist of a common α subunit and a hormone-specific β subunit (Pierce and Parsons, 1981).

Hormonal activity requires strong and specific non-covalent interactions between the α and β subunits and is conferred by the β subunit.

Secreted by the same gland and sharing low but significant sequence identity in their β subunits (~32 % in human and ~31% in rat), LH and FSH always coordinately function together in their biological activities. Their biological roles include stimulation of testicular (male) and ovarian (female) functions via steroid hormone synthesis and regulation of gametogenesis (Catt and Pierce, 1978). In female mammals, LH and FSH can induce final maturation of ovarian follicles, ovulation, estrogen and progesterone secretion, and corpus luteum formation (Figure 2). In male mammals, they can stimulate testosterone secretion and sperm production (Figure 2).

Gonadotropins, as glycoprotein hormones, are heavily glycosylated, with two N-linked oligosaccharide chains on the α subunit and either one or two (depending on species) N-linked oligosaccharide chains on the β subunit (Pierce and Parsons, 1981). There is strong evidence that glycosylation can modulate biological activities of gonadotropins. Characterizations of fully deglycosylated and partially deglycosylated gonadotropins indicated that full glycosylation is required for the full expression of bioactivities (Matzuk and Boime, 1989; Sairam, 1989). Further studies have shown that terminal carbohydrates rather than the presence of oligosaccharide chains determine the hormonal biological activities *in vivo* (Baenziger and Green, 1988; Baenziger et al., 1992).

Biological functions of reproductive hormones



Development of male sex organs (the penis, prostate gland, etc.)
Development of secondary male sex characteristics
(rooster's comb, facial hair, etc.)
Spermatogenesis

Development of female reproductive organs
(uterus, ovary, and vagina)
Development of secondary female sex characteristics
(breast, reproductive cycles, etc.)

Figure 2. Biological functions of reproductive hormones

The gonadotropins (LH and FSH) play important roles in regulation of gametogenesis and steroid hormone synthesis. The steroid hormones, including testosterone (male), estrogen (female) and progesterone (female), control the development of secondary sex characteristics in both male and female. (Taken from Randall, 1997.)

2) LH bears unique N-linked oligosaccharides

Among glycoproteins with N-linked oligosaccharides, only highly specialized populations of glycoproteins bear unique carbohydrate structures. The presence of such structures implies a special biological activity. One of the gonadotropins, LH, bears sulfated N-linked oligosaccharide chains that terminate with a sequence SO_4-4GalNAcβ1-4GlcNAcβ1-2Manα. The other gonadotropin (FSH), though synthesized within the same cells in the pituitary gland, bears sialyated N-linked glycans with terminal sequences of Siaα-Galβ1-4GlcNAcβ1-2Manα (Baenziger and Green, 1988; Green and Baenziger, 1988; Green and Baenziger, 1988). The glycosylation of LH appears to be a highly specific process, since few other glycoproteins synthesized in the pituitary gland or other places bear the same kind of N-linked glycans as LH does. (TSH is the only other pituitary hormone that is known to bear sulfated N-linked glycans (Green and Baenziger, 1988).)

Sulfated N-linked oligosaccharide chains reflect posttranslational modification in the pituitary gland. Two enzymes, GalNAc transferase and GalNAcβ1-4GlcNAcβ1-2Manα-specific sulfotransferase, have been identified to tightly regulate the synthesis of sulfated oligosaccharides of LH (Green et al., 1985; Smith and Baenziger, 1988; Smith and Baenziger, 1990). A sequence motif on LH, Pro-X-Arg (X is preferred to be a hydrophobic residue), has been identified to play a key role in the modification (Smith and Baenziger, 1992). This motif is located at 6-9 residues at the amino-terminal side of an N-linked glycosylation site on the common α subunit and the β subunit of LH, but not

on the β subunit of FSH. Recognition of the motif by GalNAc transferase switches LH into a synthesis pathway of sulfated N-linked oligosaccharides (Figure 3) (Smith and Baenziger, 1992; Smith and Baenziger, 1988). Lacking this motif, other glycoproteins (including FSH) will go through a synthesis pathway of sialyated N-linked oligosaccharides (Figure 3) (Smith and Baenziger, 1988; Stryer, 1995). (FSH, having a masked Pro-X-Arg motif on its α subunit (Smith and Baenziger, 1990), is still modified with sialyated N-linked glycans.)

3) Biological significance of sulfated oligosaccharides on LH

Baenziger and colleagues first proposed that the sulfated oligosaccharides present on LH are critical for its full biological functions (Baenziger et al., 1992) (Baenziger, 1996). In their studies, properties of native and recombinant LHs that only differed in terminal portions of their oligosaccharides were compared (Table 1). The difference in terminal sugars did not show any impact on its interactions with the LH receptor *in vitro* and *in vivo* (Baenziger et al., 1992). However, it had a significant impact on the metabolic clearance rate (MCR) of LH. Native bovine LH (bLH), bearing oligosaccharides terminating with $4SO_4$-GalNAc, is cleared from circulation at a rate more rapidly (~5 fold) than recombinant bLH that bears sialyated oligosaccharides. The terminal sulfate group appeared to be critical for controlling the MCR of LH, since either removal of the sulfate or replacement with sialic acid resulted in a significant change in the MCR. The control of the MCR is thought to be critical for the precise regulation of the circulation half-life of LH (Smith et al., 1993).

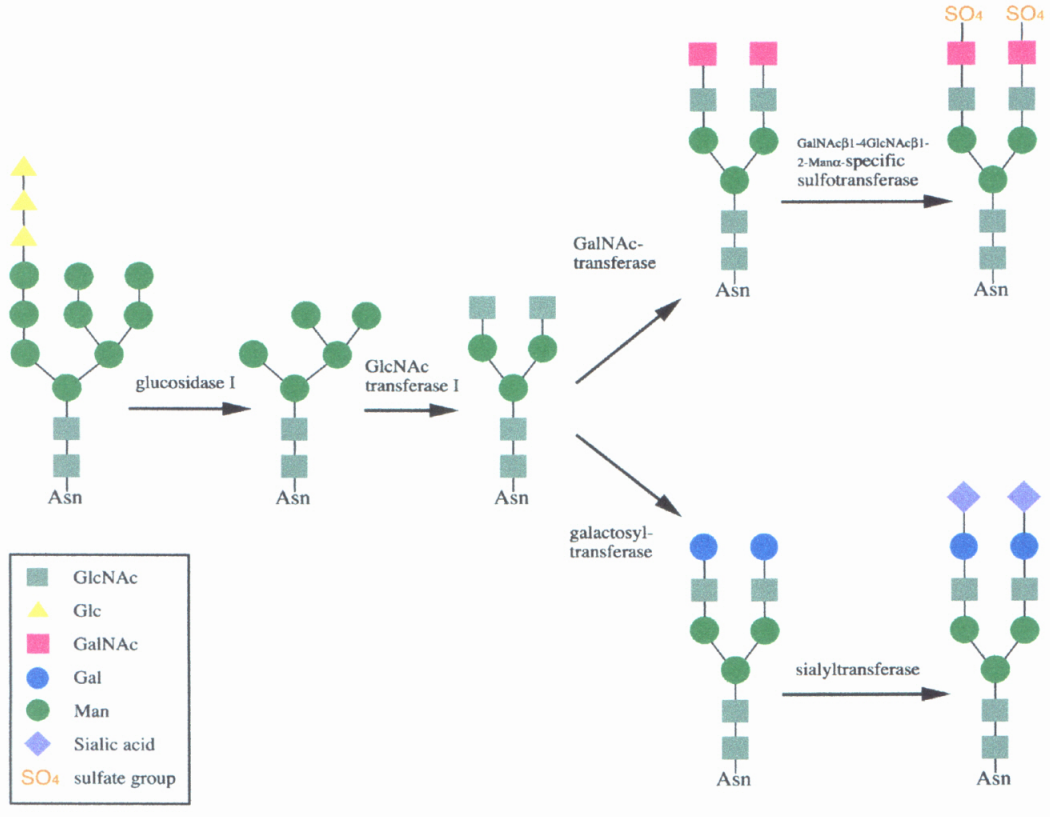


Figure 3. Pathways for biosynthesis of sulfated and sialylated N-linked oligosaccharides on pituitary glycoprotein hormones. (Modified from Smith *et al.*, 1988.)

Hormones	Terminal carbohydrate sequences	MCR (%/min)
Native bLH	SO_4-4GalNAcβ1-4GlcNAcβ1-	7.3
bLH/CHO	Siaα-Galβ1-4GlcNAcβ1-	1.7
Desulfated bLH,	GalNAcβ1-4GlcNAcβ1-	35
Desialyzed bLH/CHO	Galβ1-4GlcNAcβ1-	35

Table 1. Comparison of metabolic clearance rates (MCRs) of bovine LH with different terminal sugars

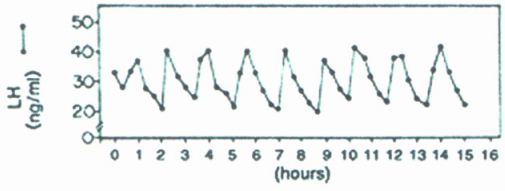
Native bovine LH (bLH) is compared with recombinant bLH expressed in Chinese hamster ovary (CHO) cells (bLH/CHO), desulfated native bLH, and desialyzed bLH/CHO. The bLH, bearing oligosaccharides terminating with $4SO_4$-GalNAc, is cleared from circulation more rapidly (4-5 fold) than bLH/CHO, which bears terminal sialic acid on its oligosaccharides. Desulfated native LH and desialyzed bLH/CHO have improperly modified oligosaccharides. Therefore, they are cleared from circulation very rapidly. (Modified from Baenziger et al, 1992.).

Blood concentrations of LH varies in a pulsatile manner throughout the menstrual cycle (Figure 4A) (Crowley and Holfer, 1985; Patton et al., 1989), resulting from secretory bursts of LH (Veldhuis et al., 1987; Veldhuis and Johnson, 1988). The characteristic intermittent mode of hormone secretion is important for hormonal functions (Desjardins, 1981; Urban et al., 1988). Precise regulation of the circulation half-life of LH is thought to be essential for generating appropriate pulsatile variations in the serum level of LH, modulating the amplitude and frequency of LH's secretory pulse, and maintaining a steady-state level of circulatory LH (Baenziger et al., 1992). Besides these functions, the regulation is also critical for rapid removal of LH after its preovulatory surge (Figure 4B), a process that prevents desensitization of the LH receptor and thereby maintains maximal hormone responsiveness *in vivo*. Overall, the sulfated oligosaccharides on LH provide a mechanism for controlling its circulatory half-life and thereby regulate its biological activities *in vivo*.

4) MR on liver endothelial cells mediates clearance of LH in the circulation

Fiete *et al.* identified a receptor within rat liver that can recognize and internalize LH (Fiete et al., 1991). Later, they demonstrated that this receptor is actually MR, since it has the same structural and antigenic properties as MR (Fiete and Baenziger, 1997) and its properties could be conferred by recombinant MR expressed in CHO cells (Fiete et al., 1997). It is obvious why MR in liver functions to uptake LH. First, the blood is the primary carrier of LH as well as other hormones. Second, liver accounts for about 25% of the total blood flow in the circulation (Patton et al., 1989), and is responsible for the

A



B

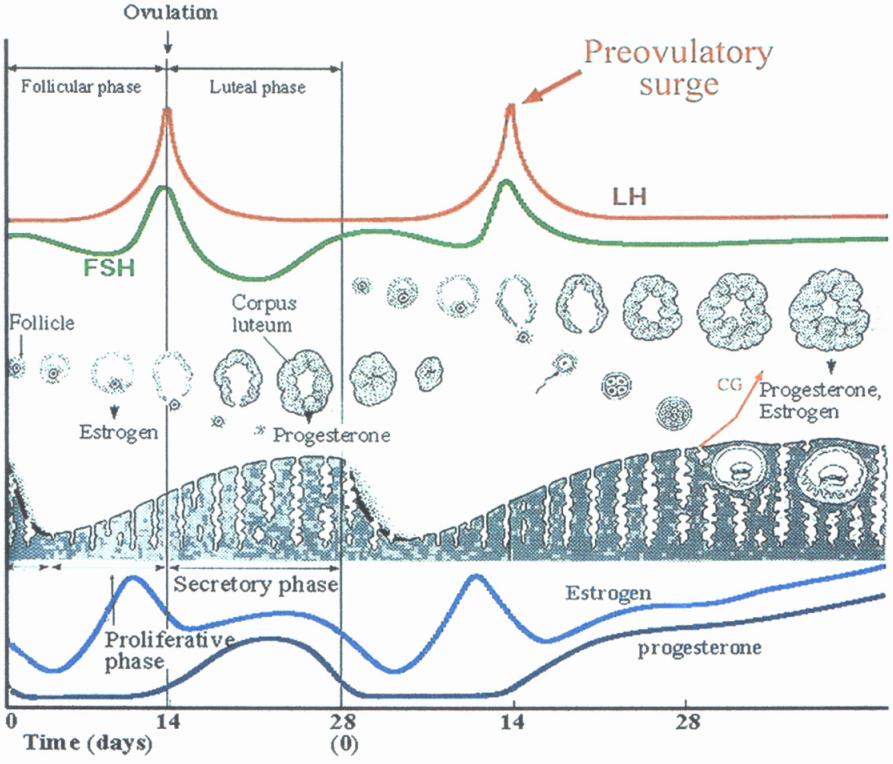


Figure 4. LH and its role in the regulation of the menstrual cycle

(A) Blood concentration of LH varies in a pulsatile manner. (From Crowley,1985.)
(B) Regulation of the menstrual cycle by periodic changes in the levels of the gonadotropins, estrogens and progesterone. Preovulatory surge of LH can trigger ovulation from one follicle and induce secretion of estrogen and progesterone. In the case of implantation (right), secretion of CG will replace LH and induce secretion of estrogen and progesterone in the first three months of pregnancy in humans. (Taken from Randall,1997.)

uptake of many hormones as well as glycoproteins. It has been shown that MR is abundantly expressed on endothelial cells and Kupffer cells in rat liver (with roughly 600,000 MRs per endothelial cell) (Fiete et al., 1991). MR is thought to bind to LH through recognition of the sulfated oligosaccharides on LH. The binding is in a pH dependent, requiring a pH above 5.5 (The tested pH range was between 4.0 and 7.5), consistent with the endocytic property of MR (Fiete et al., 1991; Simpson et al., 1999). Though MR binding to LH has a moderate binding affinity (apparent $K_D \approx 160$ nM) (Simpson et al., 1999), clearance of LH is still very efficient due to the high expression level of MR on hepatic endothelial cells.

The binding of MR to LH has been characterized, and the results showed a number of remarkable features. Studies with neoglycoproteins, consisting of BSA conjugated with SO_4-4GalNAcβ1-4GlcNAcβ1-Manα or with mannose, suggest that the mannose-specific binding site and $4SO_4$-GalNAc-specific binding site on MR are distinct (Fiete et al., 1998). As previously described in detail, the region of CRDs 4-8 accounts for mannose/fucose/GlcNAc-specific binding (Taylor et al., 1992) and the Cys-MR region for $4SO_4$-GalNAc-specific binding. Though the BSA conjugates were useful in characterizing the binding properties of MR, they could not accurately represent glycoproteins with naturally occurring oligosaccharides. The glycoprotein hormone LH often bears biantennary hybrid oligosaccharides in which one arm terminates with SO_4-4GalNAcβ1-4GlcNAcβ1-Manα and the other has terminal mannose(s) (Figure 5) (Green and Baenziger, 1988; Green and Baenziger, 1988). Studies have shown that a high

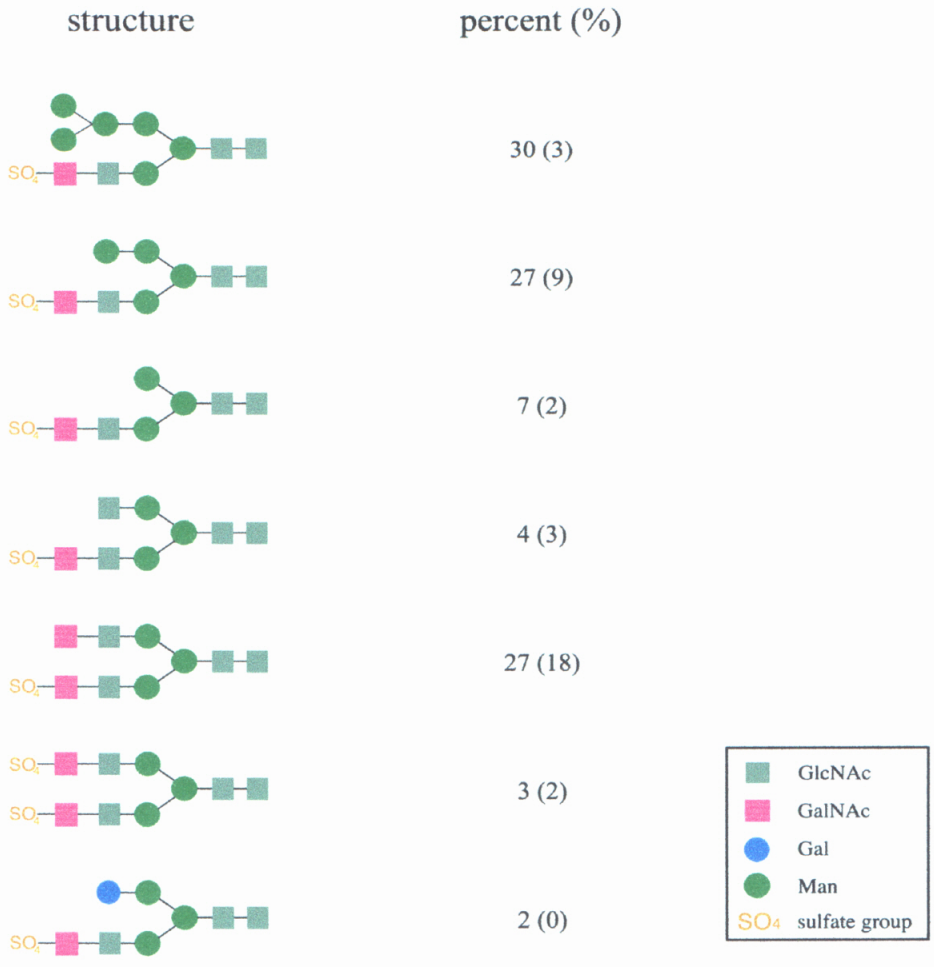


Figure 5. Structures of N-linked oligosaccharides found on LH

An analysis of sulfated oligosaccharides on LH showed that a majority of them (64%) had hybrid structures, with one or two terminal mannose residues as well as one terminal 4SO4-GalNAc residue. For each type of oligosaccharide, its percentage of the total pool of sulfated oligosaccharides is indicated, with number in parentheses indicating the percentage of structures with core fucose residues. (Taken from Green *et al.*, 1987.)

proportion of sulfated oligosaccharides on LH (about 60%) have the hybrid structures. In addition, a significant portion of oligosaccharides (37%) have fucose residues linked to the core GlcNAc, and those that do not have terminal mannoses tend to have a higher proportion of core-fucose residues (Figure 5) (Green and Baenziger, 1988; Green and Baenziger, 1988). The core fucose residues, just like terminal mannoses, could also be recognized by the region of CRDs 4-8 of MR (Stahl et al., 1978). Overall, binding of MR to LH most likely requires simultaneous recognition of the Cys-MR and the region of CRDs 4-8 to sulfated ($4SO_4$-GalNAc) and non-sulfated (mannose and fucose) carbohydrates.

LH contains three or four N-linked glycans. Although it is present at low concentrations in the blood (in the range between 3nM and 30nM) (Sherman et al., 1976), LH is still cleared rapidly at the rate of 7.3% per minute (approximately 5×10^{14} to 5×10^{15} molecules per minute). The removal process is mediated by hepatic endothelial cells and Kupffer cells, which represent about 15% and 10% of the total liver cells respectively (Arias et al., 1988) and have abundant expression of MR at the cell surface (~600,000 MR molecules per endothelial cell). *In vivo*, the binding of MR to LH is most likely achieved through avidity effects, with more than one MR binding to the N-glycans on an LH molecule, thereby achieving high binding affinity through avidity effects. The effects of avidity enhance the binding of MR to LH in two interesting ways. First, the arrangement of more than one lectin domain in a single MR molecule increases the binding affinity, from the millimolar range to the micromolar range: the K_D for CRD4

binding to α-methyl mannoside is 2.4 mM (Mullin et al., 1997), while the binding affinity of whole MR for LH is about 0.2 μM (Simpson et al., 1999). Second, with multiple MRs interacting with one LH molecule, the overall binding affinity will probably be raised from the micromolar range to nanomolar range or higher, thereby allowing MR to bind and internalize LH present at physiological concentrations (3nM to 30nM) in the blood.

Immunological roles of MR

The immune system consists of two parts, the innate immune system and the adaptive immune system. The innate immune system provides the first line of defense against many infectious microorganisms and is essential for controlling infections at an early stage, while the adaptive immune system provides a specific means of defense to clear invading pathogens and prevent subsequent re-infection with the same pathogen. MR, as a multi-functional receptor, plays important roles in both systems. In innate immunity, macrophage MR can recognize and phagocytose invading microorganisms (Martinez-Pomares and Gordon, 1999; Stahl and Ezekowitz, 1998). In adaptive immunity, MR on dendritic cells (DCs) and macrophages participates in antigen capturing and processing, which are important for presenting antigens to T lymphocytes (Martinez-Pomares and Gordon, 1999; Stahl and Ezekowitz, 1998).

The multi-lectin structure provides MR the capability to fulfill the above functions. The extracellular lectins, Cys-MR and the tandem CRD regions, are critical for

antigen capturing and antigen targeting (Martinez-Pomares and Gordon, 1999). And the transmembrane and cytoplasmic domains are crucial for the endocytic and phagocytic functions of MR (Kruskal et al., 1992). Here, I will review the biological roles of MR in the innate and adaptive immune systems in more detail.

The role of MR in innate immunity

In vertebrates, the innate immune system is essential to all individuals. It provides the first line of defense against many invading microorganisms, including bacteria, yeast, viruses and protozoa. Many of the microorganisms, in order to evade humoral immune system, have polysaccharides heavily coated on their cell walls. The molecular structures of the polysaccharides, characteristic to certain population of microorganisms, rarely exist in the host. They are classified as pathogen-associated molecular patterns (PAMPs), according to a theory proposed by Medzhitov and Janeway (Medzhitov and Janeway, 1997). PAMPs can be easily recognized by the host innate immune system via its pattern recognition receptors (PRRs), and the recognition will thereby stimulate innate immune responses.

The MR on macrophages is one of the best-characterized PRRs to date. It has been shown that MR mediates recognition and phagocytosis of many microorganisms, including bacteria, yeast, and protozoa (Stahl and Ezekowitz, 1998). On macrophages, MR-mediated recognition is specific to polysaccharide components on microorganism cell walls, such as yeast mannan, bacterial polysaccharide capsules, and some strains of

LPS and lipoarabinomannan (Ofek et al., 1995). A common feature of these polysaccharides is the structure of α-linked branched oligo-mannoses, indicating that the recognition by MR is mediated via the mannose-specific binding region - CRDs 4-8. It has been shown that MR can mediate phagocytosis independently upon ligand recognition. When non-phagocytic COS cells are transfected with human MR, they are conferred with the ability to bind and phagocytose *Candida albican* (Ezekowitz et al., 1990), *Pneumocystis carinii* (Ezekowitz et al., 1991), and *Klebsiella pneumoniae* (Kabha et al., 1995).

Expression and function of MR are modulated by many factors, including cytokines, immunoglobulin receptors, pathogens and their products (Stahl and Ezekowitz, 1998). Studies have indicated that the functions of MR are closely related to the functional state of macrophages in innate immunity. One of the best-characterized modulators is interferon-γ (IFN-γ), a TH1-type inflammatory cytokine that activates macrophages to kill intravesicular bacteria (Doyle et al., 1994; Marodi et al., 1993; Schreiber et al., 1993). Treatment of macrophages with IFN-γ enhances MR-mediated phagocytic capability, even though MR expression and MR-endocytosis are down-regulated (Marodi et al., 1993). Interestingly, the primary TH2-type cytokine interleukin-4 (IL-4), which is an antagonist for IFN-γ in adaptive immunity, acts synergistically with IFN-γ to enhance MR-mediated phagocytosis (Raveh et al., 1998). Based on these observations, it has been suggested that IFN-γ and IL-4 may cooperate to accelerate MR-mediated clearance of pathogens in innate immunity.

Upon phagocytosis of pathogens, MR can trigger a variety of responses in macrophages to facilitate intracellular killing. These responses include lysosomal enzyme secretion (Ohsumi and Lee, 1987), superoxide anion release (Berton and S., 1983), cytokine synthesis and release (Shibata et al., 1997; Yamamoto et al., 1997), and modulation of other cell surface receptors (Murai et al., 1996). Therefore, MR appears to play more than a passive role as a PRR. In conclusion, MR plays dynamic roles in response to microorganism infection, acting as a classical PRR in pathogen recognition and phagocytosis as well as a signal transduction receptor to induce intracellular killing mechanisms.

The role of MR in adaptive immunity

Pathogens can invade the body in many different ways and initiate infections anywhere. However, they eventually will be captured by specialized cells and be presented to lymphocytes to induce adaptive immune responses. Adaptive immunity occurs in the peripheral lymphoid organs - the lymph nodes, spleen and tonsils, which can efficiently trap antigens (including pathogens and their products) from the lymph and the blood. Antigens are either captured directly by antigen presenting cells (APCs) in the peripheral lymphoid organs or transported by migrating APCs that catch them at sites of infection. The captured antigens are then presented to naive T cells to initialize adaptive immunity.

Dendritic cells (DCs), macrophages and B cells are three main types of 'professional' APCs, which have the ability to present antigens *via* MHC class II molecules and deliver co-stimulatory signals to activate naive T cells. The finding that MR is expressed on DCs and macrophages suggests that MR is a key molecule in antigen recognition. Besides this function, MR also participates in antigen processing in two interesting ways.

1) MR participates in antigen processing in immature DCs

DCs are classified as immature DCs and mature DCs, according to their functional states (Banchereau and Steinman, 1998). Immature DCs capture antigens in peripheral tissues, while mature DCs, located in the peripheral lymphoid organs, function to present antigens to T lymphocytes (Banchereau and Steinman, 1998; Janeway and Travers, 1997). MR, along with other antigen capture receptors, is expressed at high levels in immature DCs (Avrameas et al., 1996) and participates in capturing antigens (Engering et al., 1997). Soluble antigens with high mannose oligosaccharides are suggested to be recognized and captured *via* the tandem CRD region of MR (Sallusto et al., 1995). They are subsequently endocytosed and delivered to endosomes and MIIC compartments, where antigens are digested by proteases and then loaded onto MHC class II molecules (Sallusto et al., 1995). Antigen uptake can trigger immature DCs to migrate to lymph nodes, where they become mature DCs and present antigens to T lymphocytes and induce adaptive immunity (Janeway and Travers, 1997) (Figure 6A).

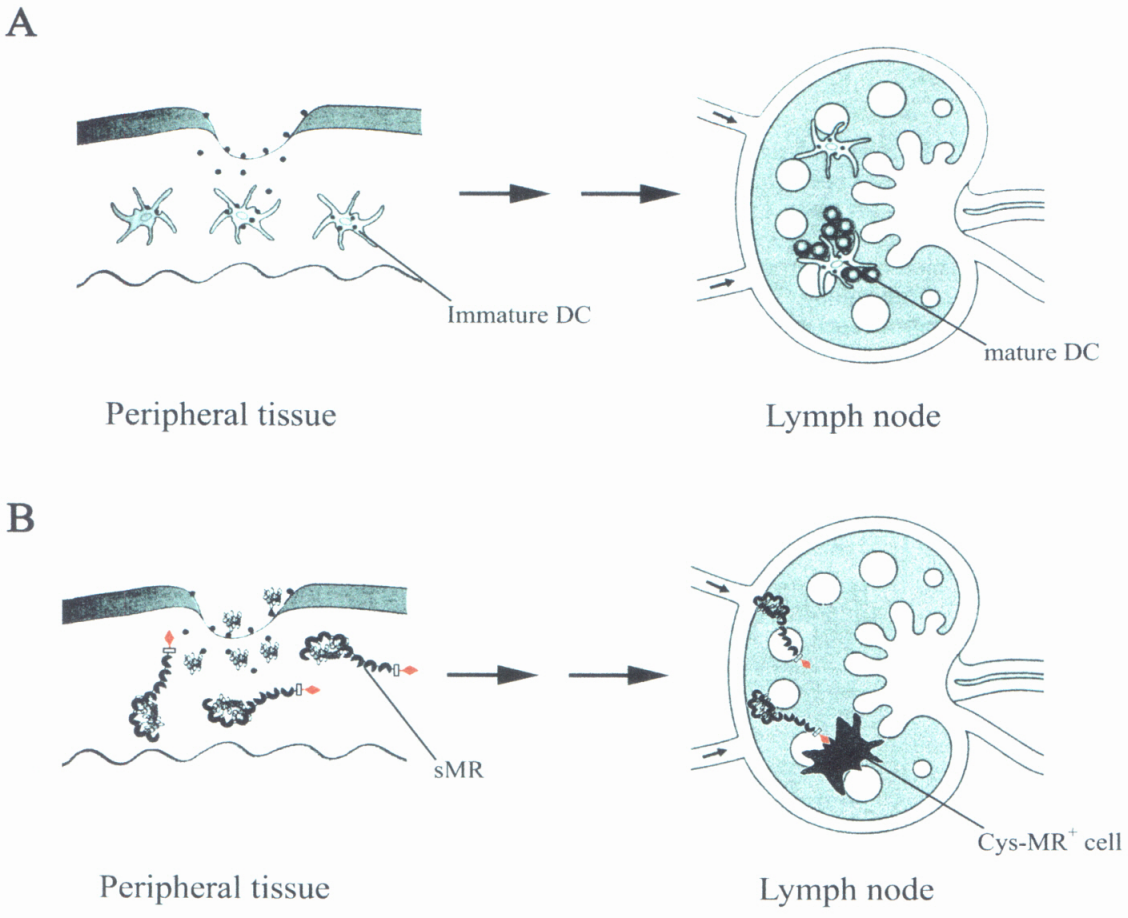


Figure 6. Antigen uptake pathways

(A) Immature dendritic cells (DCs) ingest antigens in the peripheral tissues in response to infections. Afterwards, they migrate to the peripheral lymphoid organs where they differentiate into mature DCs and present antgens to T cells. (Taken from Janeway, 1997.) (B) sMR uses its tandem CRD region to bind mannosylated antigens in the peripheral tissues, and transports them to the peripheral lymphoid organs. The Cys-MR was used to target Cys-MR$^+$ cells by which antigens are uptaken and processed. (Modified from Martinez-Pomares *et al.*, 1999.)

Interestingly, it has been speculated that MR is also involved in the antigen processing pathway of a MHC class I homolog, CD1b (Prigozy et al., 1997). It has been suggested that MR is responsible for uptake of a mycobacterial lipoarabinomannan (LAM) on monocyte-derived antigen presenting cells (a mixture of immature DCs and macrophages). The LAM is internalized and transported to late endosomes, lysosomes, and MIIC compartments. The observation that MR and CD1b molecule are co-localized in these compartments suggests that MR may deliver LAM as well as other lipoglycans to CD1b (Prigozy et al., 1997). In this way, MR may play a role in antigen processing of glycolipids for non-classical class I MHC molecules.

2) The role of MR in antigen targeting

Recently, Martinez-Pomares and Gorden proposed an alternative antigen presentation pathway mediated by MR (Martinez-Pomares and Gordon, 1999). Their proposal was mostly based on their studies on immunological ligands for Cys-MR. Using a chimeric protein with Cys-MR fused to Fc portion of human IgG1 (Cys-MR/Fc protein), Martinez-Pomares and colleagues first identified ligands of Cys-MR on specialized cell populations in the peripheral lymphoid organs of naive mice. The specialized cell populations included cells in the marginal zone of the spleen and cells in subcapsular sinus of lymph nodes (Martinez-Pomares et al., 1996). In response to a secondary immune response, the specialized cell populations were found to migrate to germinal centers in splenic white pulp and to follicular areas of lymph nodes. Interestingly, the migration showed a process similar to that of B cell follicle formation

(Martinez-Pomares et al., 1996). Two endogenous ligands for Cys-MR were identified as sialoadhesin and CD45, known as a macrophage-specific membrane molecule and leukocyte common antigen respectively (Martinez-Pomares et al., 1999). Cys-MR binding to its ligands has been shown to require the presence of sulfated carbohydrate structures on the ligands, consistent with the sulfated carbohydrate binding property of Cys-MR (Leteux et al., 2000; Martinez-Pomares et al., 1999). Although they are known as common antigens, sialoadhesin and CD45 have isoforms bearing sulfated carbohydrates that are expressed only on the specialized cell populations that are bound to Cys-MR/Fc protein (Martinez-Pomares et al., 1999). Further *in situ* studies indicated that MR and its ligands were expressed at non-overlapping sites in the peripheral lymphoid organs (Linehan et al., 1999), and the results led to a postulation that a soluble form of MR might exist and act as the counter receptor for sialoadhesin and CD45. At the same time, a soluble form of MR was identified in macrophage conditioned media and in human and mouse serum (Martinez-Pomares and Gordon, 1999; Martinez-Pomares et al., 1998). The soluble MR lacks the transmembrane and cytoplasmic domains and is probably produced by cleavage of membrane MR *via* metalloproteinases (Martinez-Pomares and Gordon, 1999). Release of soluble MR appears to be regulated by cytokines. It is up-regulated by IL-4 treatment, but down-regulated by IFN-γ treatment (Martinez-Pomares and Gordon, 1999).

An antigen uptake model mediated by soluble MR has been proposed based upon the above findings (Figure 6B). The soluble MR could act as a bifunctional molecule,

using its tandem CRD region to bind mannosylated antigens and its Cys-MR region to bind sialoadhesin or CD45. Therefore, antigens bound by soluble MR could be transported to the specialized cell populations that are bound to Cys-MR/Fc protein in the peripheral lymphoid organs and then be presented, instead of being processed by the DC-mediated pathway (Figure 6A). Along with the finding that the specialized cell populations can capture antigens and migrate to germinal centers for presentation to B cells (Martinez-Pomares et al., 1996), it has been suggested that this alternative antigen processing pathway could enhance antigen presentation required for B cell activation (Martinez-Pomares and Gordon, 1999). At present, this intriguing model needs to be further assessed *in vivo*, and further investigations on Cys-MR knock-out mice will help to evaluate the role of Cys-MR in the model (M.C. Nussenzweig, unpublished results).

Structural and biochemical characterization of Cys-MR

Given the important roles that Cys-MR plays in the process of MR recognition of its ligands, understanding the mechanism of sulfated carbohydrate recognition would be of great interest. In chapter 1, I describe the determination of the crystal structure of Cys-MR complexed with $4SO_4$-GalNAc. The structure provides the basis upon which sulfated carbohydrate recognition by Cys-MR was first defined. In chapter 2, I describe the biochemical characterization of the binding of Cys-MR to sulfated and nonsulfated carbohydrates. I also describe the determination of crystal structures of unliganded Cys-MR as well as Cys-MR complexed with $3SO_4$-Lewis^x, $3SO_4$-Lewis^a, and $6SO_4$-GalNAc.

A new mechanism of sulfated carbohydrate recognition by Cys-MR has been proposed in this chapter, and the result facilitates our understanding of the role of Cys-MR in MR binding to its sulfated carbohydrate ligands.

References:

Ancian, P., Lambeau, G., Mattei, M. G., and Lazdunski, M. (1995). The human 180-kDa receptor for secretory phospholipases A2. Molecular cloning, identification of a secreted soluble form, expression, and chromosomal localization. J. Biol. Chem. *270*, 8963-8970.

Arias, I. M., Jakoby, W. B., Popper, H., Schachter, D., and Shafritz, D. A. (1988). The liver biology and pathobioloby (Raven Press), pp. 663-704.

Ashwell, G., and Morell, A. G. (1974). The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. Adv. Enzymol. Relat. Areas. Mol. Biol. *41*, 99-128.

Avrameas, A., McIlroy, D., Hosmalin, A., Autran, B., Debre, P., Monsigny, M., Roche, A. C., and Midoux, P. (1996). Expression of a mannose/fucose membrane lectin on human dendritic cells. Eur. J. Immunol. *26*, 394-400.

Baenziger, J. U. (1996). Glycosylation: to what end for the glycoprotein hormones? Endocrinology *137*, 1520-1522.

Baenziger, J. U., and Green, E. D. (1988). Pituitary glycoprotein hormone oligosaccharides: structure, synthesis and function of the asparagine-linked oligosaccharides on lutropin, follitropin and thyrotropin. Biochim. Biophys. Acta. *947*, 287-306.

Baenziger, J. U., Kumar, S., Brodbeck, R. M., Smith, P. L., and Beranek, M. C. (1992). Circulatory half-life but not interaction with the lutropin/chorionic gonadotropin receptor is modulated by sulfation of bovine lutropin oligosaccharides. Proc. Natl. Acad. Sci. USA *89*, 334-338.

Banchereau, J., and Steinman, R. M. (1998). Dendritic cells and the control of immunity. Nature *392*, 245-252.

Berton, G., and S., G. (1983). Desensitization of macrophages to stimuli which induce secretion of superoxide anion. Down-regulation of receptors for phorbol myristate acetate. Eur. J. Immunol. *13*, 620-627.

Bijsterbosch, M. K., Donker, W., van de Bilt, H., van Weely, S., van Berkel, T. J., and Aerts, J. M. (1996). Quantitative analysis of the targeting of mannose-terminal glucocerebrosidase. Predominant uptake by liver endothelial cells. Eur. J. Biochem. *237*, 344-349.

Catt, K. J., and Pierce, J. G. (1978). Reproduction Endocrinology (Saubders).

Caux, C., Massacrier, C., Vanbervliet, B., Dubois, B., Durand, I., Cella, M., Lanzavecchia, A., and Banchereau, J. (1997). CD34+ hematopoietic progenitors from human cord blood differentiate along two independent dendritic cell pathways in response to granulocyte-macrophage colony-stimulating factor plus tumor necrosis factor alpha: II. Functional analysis. Blood *90*, 1458-1470.

Crowley, W. F., and Holfer, J. G. (1985). The Episodic Secretion of Hormones (New York: Wiley).

Desjardins, C. (1981). Endocrine signaling and male reproduction. Biol. Reprod. *24*, 1-21.

Doyle, A. G., Herbein, G., Montaner, L. J., Minty, A. J., Caput, D., Ferrara, P., and Gordon, S. (1994). Interleukin-13 alters the activation state of murine macrophages in

vitro: comparison with interleukin-4 and interferon-gamma. Eur. J. Immunol. *24*, 1441-1445.

Drickamer, K., and Taylor, M. E. (1993). Biology of animal lectins. Annu. Rev. Cell. Biol. *9*, 237-264.

Engering, A. J., Cella, M., Fluitsma, D. M., Hoefsmit, E. C., Lanzavecchia, A., and Pieters, J. (1997). Mannose receptor mediated antigen uptake and presentation in human dendritic cells. Adv. Exp. Med. Biol. *417*, 183-187.

Ezekowitz, R. A., Sastry, K., Bailly, P., and Warner, A. (1990). Molecular characterization of the human macrophage mannose receptor: demonstration of multiple carbohydrate recognition-like domains and phagocytosis of yeasts in Cos-1 cells. J. Exp. Med. *172*, 1785-1794.

Ezekowitz, R. A., Williams, D. J., Koziel, H., Armstrong, M. Y., Warner, A., Richards, F. F., and Rose, R. M. (1991). Uptake of Pneumocystis carinii mediated by the macrophage mannose receptor. Nature *351*, 155-158.

Feinberg, H., Park-Snyder, S., Kolatkar, A. R., Heise, C. T., Taylor, M. E., and Weis, W. I. (2000). Structure of a C-type Carbohydrate Recognition Domain from the Macrophage Mannose Receptor. J. Biol. Chem. *275*, 21539-21548.

Fiete, D., and Baenziger, J. U. (1997). Isolation of the SO4-4-GalNAcbeta1,4GlcNAcbeta1,2Manalpha-specific receptor from rat liver. J. Biol. Chem. *272*, 14629-14637.

Fiete, D., Beranek, M. C., and Baenziger, J. U. (1997). The macrophage/endothelial cell mannose receptor cDNA encodes a protein that binds oligosaccharides terminating with SO4-4-GalNAcbeta1,4GlcNAcbeta or Man at independent sites. Proc. Natl. Acad. Sci. USA *94*, 11254-11261.

Fiete, D., Srivastava, V., Hindsgau, O., and Baenziger, J. U. (1991). A hepatic reticuloendothelial cell receptor specific for SO4-4GalNAc beta 1,4GlcNAc beta 1,2Man alpha that mediates rapid clearance of lutropin. Cell *67*, 1103-1110.

Fiete, D. J., Beranek, M. C., and Baenziger, J. U. (1998). A cysteine-rich domain of the "mannose" receptor mediates GalNAc-4-SO4 binding. Proc. Natl. Acad. Sci. USA *95*, 2089-2093.

Goldstein, I. M. (1977). Conference on inflammation: Lysosomal hydrolases and inflammation: mechanisms of enzyme release from polymorphonuclear leukocytes. J. Endod. *3*, 329-333.

Green, E. D., and Baenziger, J. U. (1988). Asparagine-linked oligosaccharides on lutropin, follitropin, and thyrotropin. I. Structural elucidation of the sulfated and sialylated oligosaccharides on bovine, ovine, and human pituitary glycoprotein hormones. J. Bio.l Chem. *263*, 25-35.

Green, E. D., and Baenziger, J. U. (1988). Asparagine-linked oligosaccharides on lutropin, follitropin, and thyrotropin. II. Distributions of sulfated and sialylated oligosaccharides on bovine, ovine, and human pituitary glycoprotein hormones. J. Biol. Chem. *263*, 36-44.

Green, E. D., Morishima, C., Boime, I., and Baenziger, J. U. (1985). Structural requirements for sulfation of asparagine-linked oligosaccharides of lutropin. Proc. Natl. Acad. Sci. USA *82*, 7850-7854.

Ishizaki, J., Hanasaki, K., Higashino, K., Kishino, J., Kikuchi, N., Ohara, O., and Arita, H. (1994). Molecular cloning of pancreatic group I phospholipase A2 receptor. J. Biol. Chem. *269*, 5897-5904.

Janeway, C., and Travers, P. (1997). Immunobiology: The Immune System in Health and Disease, the 3rd Edition (Garland Pub).

Jiang, W., Swiggard, W. J., Heufler, C., Peng, M., Mirza, A., Steinman, R. M., and Nussenzweig, M. C. (1995). The receptor DEC-205 expressed by dendritic cells and thymic epithelial cells is involved in antigen processing. Nature *375*, 151-155.

Kabha, K., Nissimov, L., Athamna, A., Keisari, Y., Parolis, H., Parolis, L. A., Grue, R. M., Schlepper-Schafer, J., Ezekowitz, A. R., and Ohman, D. E., et al. (1995). Relationships among capsular structure, phagocytosis, and mouse virulence in Klebsiella pneumoniae. Infect. Immun. *63*, 847-852.

Kery, V., Krepinsky, J. J., Warren, C. D., Capek, P., and Stahl, P. D. (1992). Ligand recognition by purified human mannose receptor. Arch. Biochem. Biophys. *298*, 49-55.

Kruskal, B. A., Sastry, K., Warner, A. B., Mathieu, C. E., and Ezekowitz, R. A. (1992). Phagocytic chimeric receptors require both transmembrane and cytoplasmic domains from the mannose receptor. J. Exp. Med. *176*, 1673-1680.

Lambeau, G., Ancian, P., Barhanin, J., and Lazdunski, M. (1994). Cloning and expression of a membrane receptor for secretory phospholipases A2. J. Biol. Chem. *269*, 1575-1578.

Leteux, C., Chai, W., Loveless, R. W., Yuen, C. T., Uhlin-Hansen, L., Combarnous, Y., Jankovic, M., Maric, S. C., Misulovin, Z., Nussenzweig, M. C., and Feizi, T. (2000). The cysteine-rich domain of the macrophage mannose receptor is a multispecific lectin that recognizes chondroitin sulfates A and B and sulfated oligosaccharides of blood group Lewis(a) and Lewis(x) types in addition to the sulfated N-glycans of lutropin. J. Exp. Med. *191*, 1117-1126.

Linehan, S. A., Martinez-Pomares, L., Stahl, P. D., and Gordon, S. (1999). Mannose receptor and its putative ligands in normal murine lymphoid and nonlymphoid organs: In situ expression of mannose receptor by selected macrophages, endothelial cells, perivascular microglia, and mesangial cells, but not dendritic cells. J. Exp. Med. *189*, 1961-1972.

Liu, Z. H., Striker, G. E., Stetler-Stevenson, M., Fukushima, P., Patel, A., and Striker, L. J. (1996). TNF-alpha and IL-1 alpha induce mannose receptors and apoptosis in glomerular mesangial but not endothelial cells. Am. J. Physiol. *270*, C1595-1601.

Magnusson, S., and Berg, T. (1993). Endocytosis of ricin by rat liver cells in vivo and in vitro is mainly mediated by mannose receptors on sinusoidal endothelial cells. Biochem. J. *291*, 749-755.

Marodi, L., Schreiber, S., Anderson, D. C., MacDermott, R. P., Korchak, H. M., and Johnston, R. B. J. (1993). Enhancement of macrophage candidacidal activity by interferon-gamma. Increased phagocytosis, killing, and calcium signal mediated by a decreased number of mannose receptors. J. Clin. Invest. *91*, 2596-2601.

Martinez-Pomares, L., Crocker, P. R., Da Silva, R., Holmes, N., Colominas, C., Rudd, P., Dwek, R., and Gordon, S. (1999). Cell-specific glycoforms of sialoadhesin and CD45 are counter-receptors for the cysteine-rich domain of the mannose receptor. J. Biol. Chem. *274*, 35211-35218.

Martinez-Pomares, L., and Gordon, S. (1999). The Mannose Receptor and its Role in Antigen Presentation. Immunologist *7*, 119-123.

Martinez-Pomares, L., Kosco-Vilbois, M., Darley, E., Tree, P., Herren, S., Bonnefoy, J. Y., and Gordon, S. (1996). Fc chimeric protein containing the cysteine-rich domain of the murine mannose receptor binds to macrophages from splenic marginal zone and lymph node subcapsular sinus and to germinal centers. J. Exp. Med. *184*, 1927-1937.

Martinez-Pomares, L., Mahoney, J. A., Kaposzta, R., Linehan, S. A., Stahl, P. D., and Gordon, S. (1998). A functional soluble form of the murine mannose receptor is produced by macrophages in vitro and is present in mouse serum. J. Biol. Chem. *273*, 23376-23380.

Matzuk, M. M., and Boime, I. (1989). Mutagenesis and gene transfer define site-specific roles of the gonadotropin oligosaccharides. Biol. Reprod. *40*, 48-53.

Medzhitov, R., and Janeway, C. A. J. (1997). Innate immunity: impact on the adaptive immune response. Curr. Opin. Immunol. *9*, 4-9.

Mokoena, T., and Gordon, S. (1985). Human macrophage activation. Modulation of mannosyl, fucosyl receptor activity in vitro by lymphokines, gamma and alpha interferons, and dexamethasone. J. Clin. Invest. *75*, 624-631.

Mullin, N. P., Hitchen, P. G., and Taylor, M. E. (1997). Mechanism of Ca2+ and monosaccharide binding to a C-type carbohydrate-recognition domain of the macrophage mannose receptor. J Biol Chem. *272*, 5668-5681.

Murai, M., Aramaki, Y., and Tsuchiya, S. (1996). alpha 2-macroglobulin stimulation of protein tyrosine phosphorylation in macrophages via the mannose receptor for Fc gamma receptor-mediated phagocytosis activation. Immunology *89*, 436-441.

Nicolas, J. P., Lambeau, G., and Lazdunski, M. (1995). Identification of the binding domain for secretory phospholipases A2 on their M-type 180-kDa membrane receptor. J. Biol. Chem. *270*, 28869-28873.

Ofek, I., Goldhar, J., Keisari, Y., and Sharon, N. (1995). Nonopsonic phagocytosis of microorganisms. Annu. Rev. Microbiol. *49*.

Ohsumi, Y., and Lee, Y. C. (1987). Mannose-receptor ligands stimulate secretion of lysosomal enzymes from rabbit alveolar macrophages. J. Biol. Chem. *262*, 7955-7962.

Otter, M., Kuiper, J., Bos, R., Rijken, D. C., and van Berkel, T. J. (1992). Characterization of the interaction both in vitro and in vivo of tissue-type plasminogen activator (t-PA) with rat liver cells. Effects of monoclonal antibodies to t-PA. Biochem. J. *284*, 545-550.

Patton, H. D., Fuchs, A. F., Hille, B., Scher, A. M., and Steiner, R. (1989). Textbook of Physiology (W.B. Saunders company).

Pierce, J. G., and Parsons, T. F. (1981). Glycoprotein hormones: structure and function. Annu. Rev. Biochem. *50*, 465-495.

Prigozy, T. I., Sieling, P. A., Clemens, D., Stewart, P. L., Behar, S. M., Porcelli, S. A., Brenner, M. B., Modlin, R. L., and Kronenberg, M. (1997). The mannose receptor delivers lipoglycan antigens to endosomes for presentation to T cells by CD1b molecules. Immunity *6*, 187-197.

Randall, D., Burggren, W., and French, K. (1997). Animal Physiology, 4th Edition (Freeman).

Raveh, D., Kruskal, B. A., Farland, J., and Ezekowitz, R. A. (1998). Th1 and Th2 cytokines cooperate to stimulate mannose-receptor-mediated phagocytosis. J. Leukoc. Biol. *64*, 108-113.

Sairam, M. R. (1989). Role of carbohydrates in glycoprotein hormone signal transduction. FASEB J. *3*, 1915-1926.

Sallusto, F., Cella, M., Danieli, C., and Lanzavecchia, A. (1995). Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major

histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. J. Exp. Med. *182*, 389-400.

Schlesinger, P. H., Doebber, T. W., Mandell, B. F., White, R., DeSchryver, C., Rodman, J. S., Miller, M. J., and Stahl, P. (1978). Plasma clearance of glycoproteins with terminal mannose and N-acetylglucosamine by liver non-parenchymal cells. Studies with beta-glucuronidase, N-acetyl-beta-D-glucosaminidase, ribonuclease B and agalacto-orosomucoid. Biochem. J. *176*, 103-109.

Schreiber, S., Perkins, S. L., Teitelbaum, S. L., Chappel, J., Stahl, P. D., and Blum, J. S. (1993). Regulation of mouse bone marrow macrophage mannose receptor expression and activation by prostaglandin E and IFN-gamma. J. Immunol. *151*, 4973-4981.

Shepherd, V. L., and Hoidal, J. R. (1990). Clearance of neutrophil-derived myeloperoxidase by the macrophage mannose receptor. Am. J. Respir. Cell. Mol. Biol. *2*, 335-340.

Shepherd, V. L., Konish, M. G., and Stahl, P. (1985). Dexamethasone increases expression of mannose receptors and decreases extracellular lysosomal enzyme accumulation in macrophages. J. Biol. Chem. *260*, 160-164.

Shepherd, V. L., Tarnowski, B. I., and McLaughlin, B. J. (1991). Isolation and characterization of a mannose receptor from human pigment epithelium. Invest. Ophthalmol. Vis. Sci. *32*, 1779-1784.

Sherman, B. M., West, J. H., and Korenman, S. G. (1976). The menopausal transition: analysis of LH, FSH, estradiol, and progesterone concentrations during menstrual cycles of older women. J. Clin. Endocrinol. Metab. *42*, 629-636.

Shibata, N., Fukasawa, S., Kobayashi, H., Tojo, M., Yonezu, T., Ambo, A., Ohkubo, Y., and Suzuki, S. (1989). Structural analysis of phospho-D-mannan-protein complexes isolated from yeast and mold form cells of Candida albicans NIH A-207 serotype A strain. Carbohydr. Res. *187*, 239-253.

Shibata, Y., Metzger, W. J., and Myrvik, Q. N. (1997). Chitin particle-induced cell-mediated immunity is inhibited by soluble mannan: mannose receptor-mediated phagocytosis initiates IL-12 production. J. Immunol. *159*, 2462-2467.

Simpson, D. Z., Hitchen, P. G., Elmhirst, E. L., and Taylor, M. E. (1999). Multiple interactions between pituitary hormones and the mannose receptor. Biochem. J. *343*, 403-411.

Smedsrod, B., Einarsson, M., and Pertoft, H. (1988). Tissue plasminogen activator is endocytosed by mannose and galactose receptors of rat liver cells. Thromb. Haemost. *59*, 480-484.

Smith, P. L., and Baenziger, J. U. (1992). Molecular basis of recognition by the glycoprotein hormone-specific N-acetylgalactosamine-transferase. Proc. Natl. Acad. Sci. USA *89*, 329-333.

Smith, P. L., and Baenziger, J. U. (1988). A pituitary N-acetylgalactosamine transferase that specifically recognizes glycoprotein hormones. Science *242*, 930-933.

Smith, P. L., and Baenziger, J. U. (1990). Recognition by the glycoprotein hormone-specific N-acetylgalactosaminetransferase is independent of hormone native conformation. Proc. Natl. Acad. Sci. USA *87*, 7275-7279.

Smith, P. L., Bousfield, G. R., Kumar, S., Fiete, D., and Baenziger, J. U. (1993). Equine lutropin and chorionic gonadotropin bear oligosaccharides terminating with SO4-4-GalNAc and Sia alpha2,3Gal, respectively. J. Biol. Chem. *268*, 795-802.

Stahl, P., Schlesinger, P. H., Sigardson, E., Rodman, J. S., and Lee, Y. C. (1980). Receptor-mediated pinocytosis of mannose glycoconjugates by macrophages: characterization and evidence for receptor recycling. Cell *19*, 207-215.

Stahl, P. D., and Ezekowitz, R. A. (1998). The mannose receptor is a pattern recognition receptor involved in host defense. Curr. Opin. Immunol. *10*, 50-55.

Stahl, P. D., Rodman, J. S., Miller, M. J., and Schlesinger, P. H. (1978). Evidence for receptor-mediated binding of glycoproteins, glycoconjugates, and lysosomal glycosidases by alveolar macrophages. Proc. Natl. Acad. Sci. USA *75*, 1399-1403.

Stein, M., Keshav, S., Harris, N., and Gordon, S. (1992). Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. J. Exp. Med. *176*, 287-292.

Stryer, L. (1995). Biochemistry, IV Edition (Freeman).

Taylor, M. E., Bezouska, K., and Drickamer, K. (1992). Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. J. Biol. Chem. *267*, 1719-1726.

Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. D., and Drickamer, K. (1990). Primary structure of the mannose receptor contains multiple motifs resembling carbohydrate-recognition domains. J. Biol. Chem. *265*, 12156-12162.

Taylor, M. E., and Drickamer, K. (1993). Structural requirements for high affinity binding of complex ligands by the macrophage mannose receptor. J. Biol. Chem. *268*, 399-404.

Urban, R. J., Evans, W. S., Rogol, A. D., Kaiser, D. L., Johnson, M. L., and Veldhuis, J. D. (1988). Contemporary aspects of discrete peak-detection algorithms. I. The paradigm of the luteinizing hormone pulse signal in men. Endocr. Rev. *9*, 3-37.

Veldhuis, J. D., Carlson, M. L., and Johnson, M. L. (1987). The pituitary gland secretes in bursts: appraising the nature of glandular secretory impulses by simultaneous multiple-parameter deconvolution of plasma hormone concentrations. Proc. Natl. Acad. Sci. USA *84*, 7686-7690.

Veldhuis, J. D., and Johnson, M. L. (1988). A novel general biophysical model for simulating episodic endocrine gland signaling. Am. J. Physiol. *255*, E749-759.

Weis, W. I., and Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. Annu. Rev. Biochem. *65*, 441-473.

Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992). Structure of a C-type mannose-binding protein complexed with an oligosaccharide. Nature *360*, 127-134.

West, J. B. (1991). Best and Taylor's Physiological Basis of Medical Practice (Williams & Wilkins).

Wu, K., Yuan, J., and Lasky, L. A. (1996). Characterization of a novel member of the macrophage mannose receptor type C lectin family. J. Biol. Chem. *271*, 21323-21330.

Yamamoto, Y., Klein, T. W., and Friedman, H. (1997). Involvement of mannose receptor in cytokine interleukin-1beta (IL-1beta), IL-6, and granulocyte-macrophage colony-stimulating factor responses, but not in chemokine macrophage inflammatory protein 1beta (MIP-1beta), MIP-2, and KC responses, caused by attachment of Candida albicans to macrophages. Infect. Immun. *65*, 1077-1082.

Chapter 2

Crystal Structure of the Cysteine-rich Domain of Mannose Receptor Complexed with a Sulfated Carbohydrate Ligand

THE JOURNAL OF EXPERIMENTAL MEDICINE

VOLUME 191 NUMBER 7 APRIL 3, 2000

Published Semimonthly by THE ROCKEFELLER UNIVERSITY PRESS
http://www.jem.org

Crystal Structure of the Cysteine-rich Domain of Mannose Receptor Complexed with a Sulfated Carbohydrate Ligand

By Yang Liu,* Arthur J. Chirino,*‡ Ziva Misulovin,§ Christine Leteux,‖ Ten Feizi,‖ Michel C. Nussenzweig,§ and Pamela J. Bjorkman*‡

*From the *Division of Biology 156-29 and the ‡Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125; the §Department of Molecular Immunology and the Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021-6399; and the ‖Glycosciences Laboratory, Imperial College School of Medicine, Northwick Park Hospital, Harrow HA1 3UJ, United Kingdom*

Abstract

The macrophage and epithelial cell mannose receptor (MR) binds carbohydrates on foreign and host molecules. Two portions of MR recognize carbohydrates: tandemly arranged C-type lectin domains facilitate carbohydrate-dependent macrophage uptake of infectious organisms, and the NH_2-terminal cysteine-rich domain (Cys-MR) binds to sulfated glycoproteins including pituitary hormones. To elucidate the mechanism of sulfated carbohydrate recognition, we determined crystal structures of Cys-MR alone and complexed with 4-sulfated-*N*-acetylgalactosamine at 1.7 and 2.2 Å resolution, respectively. Cys-MR folds into an approximately three-fold symmetric β-trefoil shape resembling fibroblast growth factor. The sulfate portions of 4-sulfated-*N*-acetylgalactosamine and an unidentified ligand found in the native crystals bind in a neutral pocket in the third lobe. We use the structures to rationalize the carbohydrate binding specificities of Cys-MR and compare the recognition properties of Cys-MR with other β-trefoil proteins.

Key words: β-trefoil protein • hydrogen bond network • multilectin receptor • pituitary hormones • sulfated GalNAc

Introduction

The innate immune response is critical for detection and elimination of infectious microorganisms. Carbohydrates on microbial cell walls are common targets for pattern recognition receptors on immune cells such as macrophages, which recognize a variety of antigens (1). The mannose receptor (MR),[1] expressed on some macrophages, epithelial, and endothial cells (2, 3), is the prototype member of a family of multilectin receptors that recognize carbohydrates on the cell walls of infectious organisms (4). Carbohydrate recognition by MR facilitates macrophage uptake of bacteria, yeast, and parasites, thereby contributing to innate immunity towards these pathogens. MR is also thought to play a role in the adaptive immune response by transporting antigens to MHC class II–containing compartments in immature dendritic cells for antigen processing and presentation to T cells (5), and by delivering mycobacterially derived glycolipids to endosomes for T cell presentation by the MHC class I–like molecule CD1b (6).

MR is a 180-kD type I transmembrane protein that contains three extracellular regions: an NH_2-terminal cysteine-rich domain (Cys-MR), a domain containing fibronectin type II repeats, and eight tandem C-type lectin carbohydrate-recognition domains (CRDs; 4). The extracellular domains are linked to a transmembrane region and a small cytoplasmic domain. These structural characteristics are shared among members of the MR family, which includes the phospholipase A2 receptor (7–9), the dendritic cell receptor DEC-205 (10), and an MR-like receptor expressed in epithelium (11). The CRDs of the extracellular region mediate calcium-dependent binding to sugars such as mannose, fucose, and *N*-acetylglucosamine that are commonly found on microorganisms, but rarely seen in sufficient density in terminal positions of mammalian oligosaccharides

Address correspondence to Pamela J. Bjorkman, Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125. Phone: 626-395-8350; Fax: 626-792-3683; E-mail: bjorkman@cco.caltech.edu

[1]*Abbreviations used in this paper:* 4-SO_4-GalNAc, 4-sulfated-*N*-acetylgalactosamine; aFGF, acidic FGF; bFGF, basic FGF; CRD, carbohydrate-recognition domain; Cys-MR, cysteine-rich domain of the MR; FGF, fibroblast growth factor; MIRAS, multiple isomorphous replacement including anomalous scattering; MR, mannose receptor; SBP, sulfate binding protein.

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/2000/04/1105/11 $5.00
Volume 191, Number 7, April 3, 2000 1105–1115
http://www.jem.org/cgi/current/full/191/7/1105

(12, 13). Three-dimensional structures of members of the C-type lectin superfamily (14) have facilitated mechanistic studies of sugar binding by the MR CRDs (15), and the crystal structure of the principal mannose-recognition domain of MR (CRD4) has now been determined.[2]

Recent studies demonstrate that the Cys-MR region also has carbohydrate-binding properties. Cys-MR binds to glycoproteins containing sulfated sugars that terminate in SO_4-4GalNAcβ1,4GlcNAcβ1-, such as the sulfated hormones lutropin and thyroid stimulating hormone (16–18). Further studies of its carbohydrate-recognition properties indicate that Cys-MR can also bind chondroitin sulfates A and B and sulfated oligosaccharides of the Lewis^a and Lewis^x types (19). MR binding of hormones is thought to regulate their bioavailability (20) and facilitate rapid clearance of lutropin from the serum through binding to MR on hepatic endothelial cells (21). Studies using an Fc chimeric protein containing mouse Cys-MR revealed the presence of additional ligand(s) for Cys-MR. The fusion protein binds to cells in germinal centers and to macrophages from the splenic marginal zone and lymph node subcapsular sinus (22). The Cys-MR interaction with cells in the spleen results from the binding of sulfated carbohydrates (19). These results suggest that Cys-MR directs MR-bearing cells toward germinal centers during immune responses, and that sulfated carbohydrate ligands regulate the trafficking and function of cells bearing MR.

To establish the mechanism of sulfated sugar recognition by Cys-MR, we solved crystal structures of Cys-MR alone and complexed with the monosaccharide 4-sulfated-*N*-acetylgalactosamine (4-SO_4-GalNAc). The structures of Cys-MR and the principal mannose-binding domain of MR (CRD4)[2] provide a framework for modeling the structures of domains in the MR family of proteins, and facilitate an understanding of the role of MR in binding carbohydrates in antimicrobial immunity and in hormone regulation.

Materials and Methods

Expression and Purification of Soluble Cys-MR. A Cys-MR-Fc fusion protein (hydrophobic signal sequence and the first 139 residues of mouse MR, a 2-residue linker, and the Fc region of human IgG1) was produced in 293T cells (22). Supernatants from transfected cells were harvested 13 d after transfection. The Cys-MR-Fc protein was collected on protein A beads (Pierce Chemical Co.), eluted with elution buffer (Pierce Chemical Co.), dialyzed into PBS, and concentrated to 5 mg/ml using a centricon filter (Amicon). The chimeric protein was digested with papain coupled to agarose beads (0.4 U/mg of Cys-MR-Fc; Sigma Chemical Co.) in PBS, 10 mM EDTA, 10 mM L-cysteine at 37°C for 1 h. The reaction was terminated by adding iodoacetamide to a concentration of 30 mM and centrifuging out the papain beads. Fc fragments and undigested Cys-MR-Fc were removed using protein A beads. The sequence derived from NH_2-terminal sequencing of purified Cys-MR begins 19 residues after the initiating methionine (NH_2-terminal sequence of mature protein is Leu-Asp-Ala-Arg-Gln; data not shown). Purified human MR starts 18 residues after the initiating methionine with the sequence Leu-Leu-Asp-Ala-Arg-Gln (23), indicating that the position at which the hydrophobic leader peptide is cleaved in the human and mouse versions of MR differs by one residue. The numbering scheme used in this paper corresponds to the numbering of mature mouse Cys-MR.

Crystallization and Data Collection. Crystals (space group $P2_12_12_1$; a = 39.6 Å, b = 41.3 Å, c = 98.8 Å; one molecule per asymmetric unit) were grown at 4°C in 1:1 hanging drops containing Cys-MR (40 mg/ml in 10 mM Hepes, pH 7.4, 10 mM NaCl) and 30% (wt/vol) PEG 8000, 0.2 M $(NH_4)_2SO_4$. Before data collection, crystals were transferred to a cryoprotectant solution (10 mM Hepes, pH 7.0, 32% PEG 8000, 0.2 M $[NH_4]_2SO_4$, 10% glycerol). A native data set used for multiple isomorphous replacement including anomalous scattering (MIRAS) phase determination (Native I; Table I) was collected at −150°C from a single crystal to 1.9 Å using an R-AXIS II detector (Molecular Structures Corp.) mounted on a rotating anode generator (model R200; Rigaku). Heavy atom derivative data sets (prepared by soaking native crystals in synthetic mother liquor containing heavy atoms for 7 d) and the 1.7 Å native data set used for refinement (Native II; Table I) were collected using the same procedures. Crystals in complex with 4-SO_4-GalNAc (space group $P2_12_12_1$; a = 39.4 Å, b = 41.5 Å, c = 100.3 Å; one molecule/asymmetric unit) were obtained by soaking Cys-MR crystals in 2 mM 4-SO_4-GalNAc (Sigma Chemical Co.), 32% PEG 8000, 0.2 M $(NH_4)_2SO_4$ for 10 d at 4°C. Native data were collected to 2.2 Å as described above. Data were processed and scaled with the HKL package (24; Table I).

Structure Determination and Refinement. Heavy atom positions were identified from difference Patterson (mercury derivative) or difference Fourier (platinum derivative) maps. Two mercury and three platinum sites were refined and phases were calculated using SHARP (25). An initial MIRAS electron density map with a mean figure of merit of 0.56 was calculated to 2.5 Å resolution and improved by solvent flattening using Solomon (26). The experimental map was of high quality, with no main chain breaks and clearly identifiable density for most side chains and the unidentified ligand. A skeleton of the map (27) was used as a starting point for model building using the program O (28). Model refinement was performed with CNS (29) using a bulk solvent correction, refinement of individual temperature (B) factors, the maximum likelihood target function, and protocols to minimize R_{free} (30). In later rounds of refinement, water molecules were built into peaks >3.0 σ in difference Fourier maps. The refined model (R_{free} = 20.8%; R_{cryst} = 19.8%) includes residues 2–135 of Cys-MR, 264 water molecules, 1 sulfate group, and no residues in disallowed regions of the Ramachandran plot. There is no ordered density for residues 1 and residues 136–139 of Cys-MR, or for the two-residue linker inserted before the Fc. Ordered density is not observed in the vicinity of the potential N-linked glycosylation site at Asn87.

The structure of Cys-MR complexed with 4-SO_4-GalNAc was solved at 2.2 Å resolution by molecular replacement using native Cys-MR as a search model. The model was refined using CNS (29) as described above. An initial model of 4-SO_4-GalNAc was obtained by combining coordinates for GalNAc and a sulfate group, ensuring proper connectivity and bond angles. The ligand was modeled into a large continuous electron density feature with the sulfate group positioned identically to the sulfate in the native structure. The 4-SO_4-GalNAc molecule is well de-

[2]Feinberg, H., S. Park-Snyder, A.R. Kolatkar, C.T. Heise, M.E. Taylor, and W.I. Weis, manuscript submitted for publication.

Table I. *Data Collection, Phasing, and Refinement Statistics*

Data set	Resolution	No. of observations	Unique reflections	Percent complete*	R_{merge}†	$I/\sigma I$	Phasing power§
	Å				%		
Native I	25.0–1.89	104,666	13,382	98.5	3.8	23.4	
	(1.96–1.89)			(90.2)	(13.3)	(6.6)	
Native II	25.0–1.70	152,280	18,536	98.1	5.5	21.0	
	(1.77–1.70)			(91.3)	(20.5)	(5.9)	
4-SO_4-GalNAc	25.0–2.2	86,630	8,984	97.5	4.5	24.7	
Complex	(2.28–2.20)			(95.5)	(10.5)	(11.8)	
5.0 mM	25.0–2.7	49,974	4,822	93.3	9.0	15.4	2.67
K_2PtCl_4	(2.79–2.70)			(78.4)	(26.8)	(3.7)	
5.0 mM	25.0–2.1	67,012	9,875	98.0	5.3	23.7	1.49
EMP	(2.17–2.10)			(88.7)	(19.0)	(5.7)	

Refinement	Native	4-SO_4-GalNAc complex
Resolution (Å)	25.0–1.7	25.0–2.2
Reflections in working set $\|F\|>0$	17,054	8,156
Reflections in test set $\|F\|>0$	1,078	483
R_{cryst}/R_{free} (%)‖	19.8/20.8	21.3/23.2
Rmsd bond length (Å)	0.008	0.006
Rmsd angles (deg)	1.89	1.49
No. of atoms		
Protein	1,085	1,085
Water	264	153
Ligand	5	19
Ramachandran plot		
Most favored region (%)	86.1	82.8
Additional allowed region (%)	13.9	16.4
Generously allowed region (%)	0.0	0.8
Disallowed region (%)	0.0	0.0

Values in parentheses indicate the high-resolution shells. EMP, ethyl mercury phosphate.
*Complete, no. of independent reflections/total theoretical number.
†R_{merge} $(I) = [\Sigma|I(i) - <I(h)>|/\Sigma I(i)]$, where $I(i)$ is the ith observation of the intensity of the hkl reflection, and $<I>$ is the mean intensity from multiple measurements of the hkl reflection.
§Phasing power for acentric reflections: rms f_h/E, where f_h is the heavy-atom structure factor amplitude, and E is the residual lack of closure error.
‖R_{cryst} $(F) = \Sigma_h||F_{obs}(h)| - |F_{calc}(h)||/\Sigma_h|F_{obs}(h)|$, where $|F_{obs}(h)|$ and $|F_{calc}(h)|$ are the observed and calculated structure factor amplitudes for the hkl reflection. R_{free} is calculated over reflections in a test set not included in atomic refinement.

fined in Fo-Fc maps contoured at 3.0 σ and 2Fo-Fc maps contoured at 1.1 σ, with the exception of the *N*-acetyl group, which has weak electron density. The refined model (R_{free} = 23.2%; R_{cryst} = 21.3%) includes 134 amino acids (residues 2–135 of Cys-MR), 153 water molecules, and the 4-SO_4-GalNAc group. None of the residues are in disallowed regions of the Ramachandran plot.

Fig. 1, A and B, Fig. 2, A–D, and Fig. 3, A and B, were made using MOLSCRIPT (31) and Raster3D (32). Electrostatic potentials were calculated, and Fig. 2 E was made using GRASP (33).

Results

Structure Description. Mouse Cys-MR was expressed as an Fc fusion protein and purified from the supernatants of transfected 293T cells. The Cys-MR crystal structure was determined by MIRAS and refined to 1.7 Å resolution (Table I). Cys-MR is a globular protein of approximate dimensions 45 × 40 × 35 Å, which consists of 12 antiparallel β strands arranged in a pattern with approximate threefold internal symmetry (Fig. 1, A and B). Strands 1, 4, 5, 8, 9, and 12 form a centrally-located six-stranded antiparallel β

barrel covered by the remaining β strands. This folding topology, known as a β-trefoil (34), was originally described for the structure of soybean trypsin inhibitor (35), and is found in other proteins such as the ricin B chain (36), fibroblast growth factors (FGFs [37]), IL-1α and β (38, 39), and the actin binding protein hisactophilin (40). Members of this family generally share little or no detectable sequence similarity (Fig. 1 C), but weak sequence similarity between Cys-MR and the ricin B chain was previously reported (41), thus anticipating the structural similarity. Following the conventions for the description of the fold (35, 34), Cys-MR contains three lobes (I, β strands 1–4; II, β strands 5–8; and III, β strands 9–12; Fig. 1, A and B) that are related by approximate three-fold internal symmetry.

The six cysteines for which Cys-MR is named are involved in three disulfide bonds (Fig. 1, A–C). The first, between Cys16 and Cys30, connects β strand 2 and the loop between strands 3 and 4 in lobe I. The second disulfide bond, located in lobe II, is related to the first by the pseudo three-fold symmetry, connecting Cys55 in strand 6 to Cys72 in the loop between strands 7 and 8 (Fig. 1 B). Two disulfide bonds are found in similar positions in the structure of the ricin B chain (Fig. 1, A and C), which allowed correct prediction of the observed disulfide bonding pattern of Cys-MR (41). The third disulfide bond in Cys-MR is not related to the others by the internal three-fold symmetry; instead, it connects Cys83 in strand 8 to the Cys130 in the extended polypeptide chain COOH-terminal to strand 12. This disulfide bond anchors the extended region after strand 12, a feature not found in β-trefoil proteins such as ricin and FGF (Fig. 1 C), to the remainder of the Cys-MR structure. In lobe III, the residues at the positions analogous to the cysteines in lobes I and II are Tyr97 in strand 10 and Gly113 in the loop between strands 11 and 12, whose Cα atoms are separated by 7.1 Å (compared with 5.0 and 5.1 Å for the Cα separation between the cysteines involved in the disulfide bonds in the first two lobes). Even if residues 97 and 113 were substituted by cysteines, a disulfide bond could not form because the Cα atom of residue 113 faces in the opposite direction to the side chain of residue 97 (Fig. 1 B). The reorientation of residue 113 compared with its lobe I and II counterparts occurs because the loop between strands 11 and 12 is shorter than the counterpart loops in lobes I and II (Fig. 1, B and C). Lengthening the loop between strands 11 and 12 would likely disturb ligand binding by Cys-MR, as this

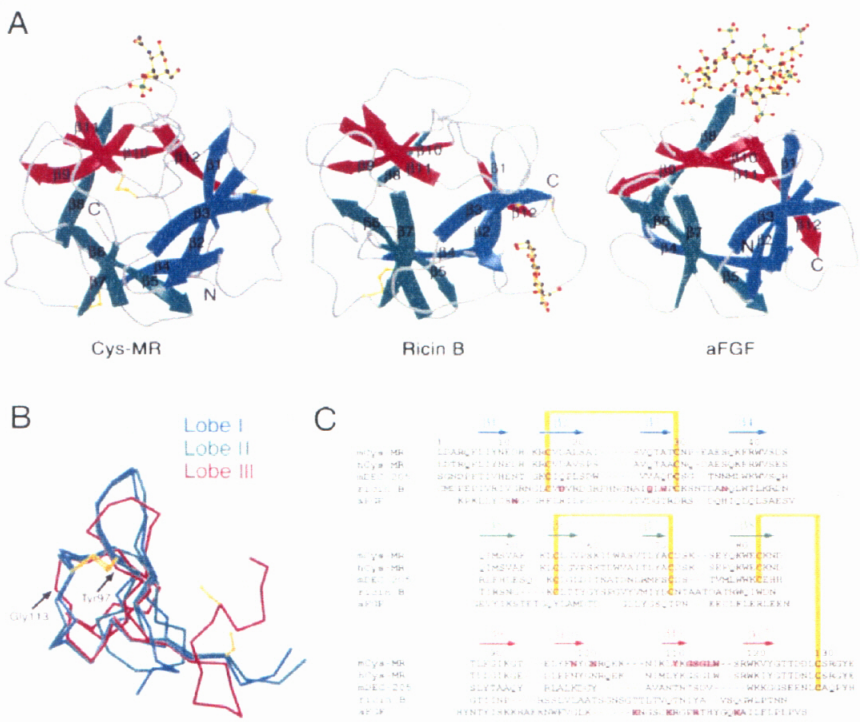


Figure 1. Comparisons between Cys-MR and other β-trefoil proteins. (A) Ribbon diagrams comparing the crystal structures of Cys-MR, a portion of the ricin B chain (residues 1–136 with N-linked carbohydrates omitted; PDB code 2AAI [reference 36]), and human aFGF (PDB code 2AXM [reference 45]). N and C-termini are labeled, disulfide bonds are yellow, and lobes I, II, and III are indicated in different colors. Each structure is depicted with a bound ligand (4-SO_4-GalNAc for Cys-MR, galactose for ricin B chain, and sulfated heparin decasaccharide for aFGF). Rms deviations for Cys-MR superimposed with ricin B chain: 1.6 Å (106 Cαs); Cys-MR superimposed with FGF: 1.9 Å (97 Cαs); ricin B chain superimposed with FGF: 1.6 Å (101 Cαs). (B) Cα superposition of Cys-MR lobes I, II, III based on Cα atoms in β strands. (C) Sequence alignment of Cys-MR and related proteins. Numbering refers to mouse Cys-MR only (see Materials and Methods). The sequences of mouse (m) and human (h) Cys-MR are aligned with mouse DEC-205, human aFGF, and the β-trefoil region of the ricin B chain. Sequences were aligned using ALIGN (available at http://www2.igh.cnrs.fr//bin/align-guess.cgi) in the case of Cys-MR and DEC-205, or based upon structural information for the sequences with known crystal structures, resulting in the following percent amino acid identities: mouse and human Cys-MR: 84%; mouse Cys-MR and DEC-205: 20%; Cys-MR and ricin B chain: 18%; Cys-MR and FGF: 24%; ricin B chain and FGF: 17%. The approximate locations of β strands 1–12 as derived from the structure of mouse Cys-MR are indicated by arrows above the sequences. Yellow lines indicate cysteine residues involved in disulfide bonds in Cys-MR. Ligand-binding residues identified from the crystal structures of Cys-MR bound to 4-SO_4-GalNAc, ricin B bound to galactose (reference 36), and aFGF bound to sulfated heparin decasaccharide are indicated by colored boxes (reference 45).

loop forms the side of the ligand binding pocket (see below).

The Cys-MR structure includes an as yet unidentified ligand that cocrystalizes with the protein. Electron density that cannot be accounted for by the amino acid sequence, N-linked carbohydrate, or water molecules is found in lobe III near the loop connecting strands 11 and 12 (Fig. 2 A, and Table II). Components of the crystallization buffer, which includes ammonium sulfate and Hepes (*N*-[2-hydroxyethyl] piperazine-*N*′-[2-ethanesulfonic acid]; see Materials and Methods), are possible sources of the ligand. A sulfate group fits into the tetrahedral portion of the electron density, and an anomalous difference Fourier map calculated using the native data and refined phases revealed a 7.0 σ peak at the presumed sulfur atom location (Fig. 2 A). The unidentified ligand is unlikely to be ammonium sulfate since the remaining portion of the density is too large to represent an ammonium group and ammonium sulfate is expected to be completely ionized under the near-neutral pH conditions of the crystallization. A molecule of Hepes can be fit into the electron density with its SO_3 group occupying the position of the sulfate group, but there is no density to account for half of the piperazine ring or the hydroxyethyl group. In the absence of conclusive identification of the ligand, it was modeled as a sulfate group in the refined native Cys-MR structure (Table II).

Crystal Structure of Cys-MR Bound to 4-SO_4-GalNAc. To identify the binding site for sulfated carbohydrates on Cys-MR and elucidate the binding interactions, we solved the structure of Cys-MR complexed with 4-SO_4-GalNAc, which was identified as a ligand by its ability to inhibit Cys-MR binding to lutropin (19). The complex structure was determined at 2.2 Å by molecular replacement using data collected from crystals soaked with the ligand. Difference

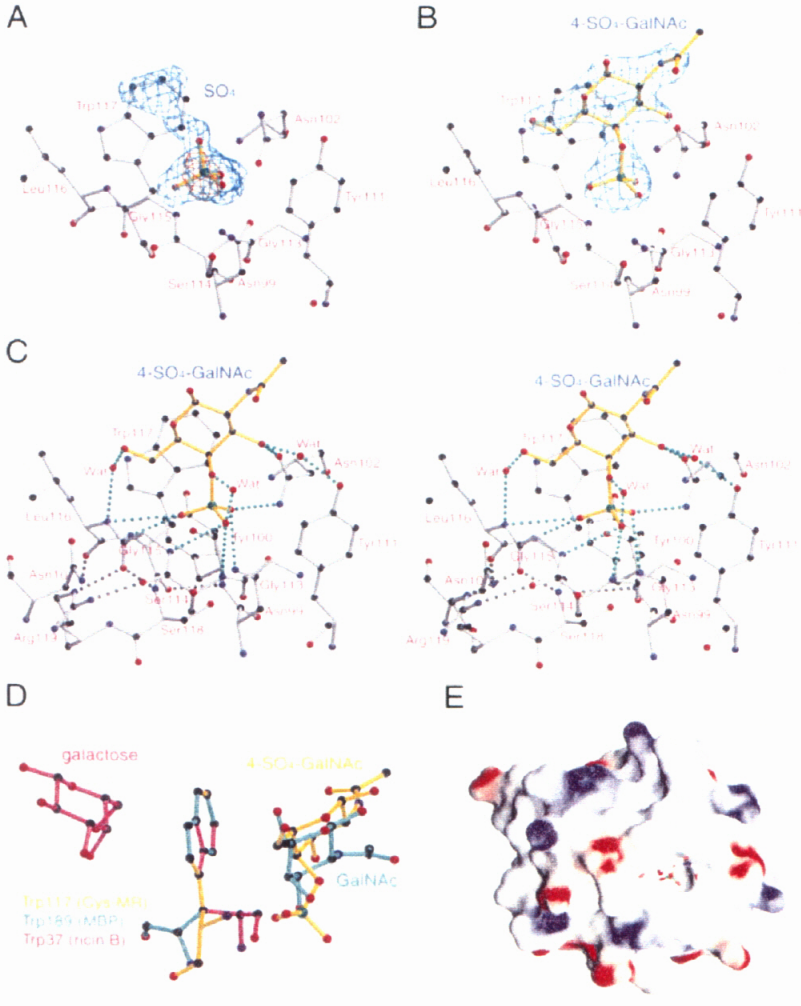


Figure 2. The Cys-MR interaction with ligands. (A) Cys-MR residues in the vicinity of the unidentified ligand (modeled as a sulfate group) derived from the refined model of native Cys-MR. Electron density for the unidentified ligand is indicated in cyan (Fo-Fc map contoured at 3.0 σ) and red (anomalous difference Fourier map contoured at 3.0 σ calculated using native data and refined phases). The peak corresponding to the sulfur of the unidentified ligand (7.5 σ) was the fourth highest peak in the native anomalous difference Fourier. Other peaks are located at the sulfur atoms of the six cysteines and of Met48 (5.5–9.5 σ). (B) Cys-MR residues in the vicinity of 4-SO_4-GalNAc derived from the refined model of the Cys-MR/4-SO_4-GalNAc complex structure. The 4-SO_4-GalNAc ligand is shown superimposed on electron density derived from an Fo-Fc annealed omit map (reference 51) contoured at 3.0 σ. (C) Stereo view of the interaction between 4-SO_4-GalNAc and Cys-MR. Hydrogen bonds between ligand and protein atoms (see Table II) are indicated by dotted green lines. Hydrogen bonds linking the CO groups of residues in which the NH group donates to a sulfate oxygen atom (see text) are indicated by dotted black lines. (D) Comparison of the stacking interactions between tryptophan and galactose rings from the complexed structures of Cys-MR (yellow bonds), the galactose binding mutant of mannose-binding protein (reference 52; PDB code 1AFB, green bonds), and ricin B chain (reference 36; PDB code 2AAI, purple bonds). (E) 4-SO_4-GalNAc is shown on the molecular surface of Cys-MR with colors highlighting the electrostatic potential calculated by GRASP (reference 33). Electrostatic potential is plotted from −10 kT/e (electronegative; red) to +10 kT/e (electropositive; blue), with white indicating electroneutrality.

Table II. *4-SO_4-GalNAc Interactions with Cys-MR*

Between sulfate and Cys-MR					Between GalNAC and Cys-MR			
Bond type	Ligand atom	Protein atom	Distance Complex	Distance Native	Bond type	Ligand atom	Protein atom	Distance
			Å	Å				Å
H bond	OS1	Asn102 ND2	3.1	3.2	H bond	3-OH	Asn102 OD1	2.6
H bond	OS1	Asn99 ND2	3.2	3.0	H bond	6-OH	Leu116 NH	3.6
H bond	OS2	Ser114 NH	3.5	3.4	H bond	3-OH	Wat-Tyr111 OH	2.9/2.7
H bond	OS2	Gly115 NH	3.2	3.1	H bond	6-OH	Wat-Leu116 NH	3.0/3.2
H bond	OS3	Leu116 NH	3.4	3.3	VDW	C3	Asn102 OD1	3.4
H bond	OS3	Trp117 NH	3.0	2.9	VDW	C3	Trp117 CZ3	3.9
VDW	OS1	Tyr111 CE2	3.5	3.3	VDW	C3	Trp117 CE3	4.0
VDW	OS1	Trp117 CB	3.5	3.7	VDW	C4	Trp117 CD2	3.8
VDW	OS2	Gly115 C	3.2	3.2	VDW	C4	Trp117 CE3	3.4
VDW	OS2	Gly113 CA	3.2	3.4	VDW	C4	Trp117 CZ3	3.8
VDW	OS3	Gly115 C	3.2	3.1	VDW	C5	Trp117 CD2	3.7
VDW	OS3	Gly115 CA	3.1	3.2	VDW	C5	Trp117 CE2	3.9
VDW	OS3	Trp117 CB	3.3	3.3	VDW	C5	Trp117 CE3	3.8
					VDW	C6	Trp117 CG	3.4
					VDW	C6	Trp117 CD1	3.6
			Distance	Distance	VDW	C6	Trp117 CD2	3.4
Bond type	Ligand atom	Ligand atom	Complex	Native	VDW	C6	Trp117 CE2	3.6
H bond	OS2	Wat-OS4	2.8/3.2	2.8/3.4	VDW	C6	Leu116 CD1	3.6
					VDW	6-OH	Leu116 CB	3.4

The following distance and geometry criteria were used for assiging hydrogen bonds: a distance of < 3.5Å, and a hydrogen bond angle of $>90°$. The maximum distance allowed for a van der Waals interaction was 4.0 Å. H bond, hydrogen bond; VDW, van der Waals interaction; Wat, water molecule.

maps reveal unambiguous electron density for the entire ligand, with the sulfate group positioned in the same location as the sulfate of the unidentified ligand found in the native structure (Fig. 2 B, and Table II). The protein portions of both structures are nearly identical (0.34 Å rms deviation for Cα atoms; 0.54 Å rms deviation for all atoms); thus, there are no major main chain or side chain rearrangements induced by ligand binding. Since the Cys-MR binding site is occupied in the native structure, we cannot address whether conformational changes occur upon ligand binding. However, ligand-induced conformational changes have not been observed in other lectin structures, which, in their unoccupied states, generally contain preformed binding sites in which ordered water molecules contribute hydrogen bonds that substitute for hydrogen bonds from the sugar hydroxyls (42).

The 4-SO_4-GalNAc ligand binds in a pocket formed on one side by residues 111–116 in the loop between strands 11 and 12, on the other side by the side chains of residues including Asn102 and Trp117, and on the bottom by the Asn99 side chain. The sulfate portion of 4-SO_4-GalNAc points into the pocket, making extensive hydrogen bonds and van der Waals contacts with residues in the 11–12 loop region (Fig. 2 C). The corresponding region of DEC-205 shares little sequence similarity with Cys-MR and contains a deletion (Fig. 1 C), thus rationalizing the observation that the cysteine-rich domain of DEC-205 does not bind lutropin (19).

The galactose ring of 4-SO_4-GalNAc, which is in the chair conformation, is positioned with the polar A face exposed to solvent and the nonpolar B face stacked against Trp117, making numerous van der Waals interactions with Asn102, Leu116, and Trp117 (Table II, and Fig. 2 D). The stacking interaction between the galactose ring and Trp117 is reminiscent of interactions seen in the structures of other protein–galactose complexes, including plant lectins, bacterial toxins, galectins, and the galactose-binding mutant of mannose-binding protein, in which the nonpolar face of the galactose moiety stacks against the tryptophan or phenylalanine side chains (42; Fig. 2 D). In addition to van der Waals interactions, the 3- and 6-hydroxyl groups of the galactose ring make hydrogen bonds with Asn102 and Leu116, and participate in two water-mediated hydrogen bonds with the protein (Fig. 2 C, and Table II). The *N*-acetyl group makes no contacts with the protein.

The specificity of Cys-MR for sulfated carbohydrates is explained at the structural level by the multitude of interactions between the protein and the sulfate group of 4-SO_4-GalNAc. Of the eight direct hydrogen bonds between Cys-

MR and the ligand, six involve the sulfate group (Table II). The first sulfate oxygen accepts two hydrogen bonds, one from the Asn99 side chain and the other from the Asn102 side chain. The other free oxygens of the sulfate group accept four hydrogen bonds, all involving main chain NH groups of the protein – the NH groups of Ser114 and Gly115 with the sulfate OS2, and the NH groups of Leu116 and Trp117 with the sulfate OS3. The NH groups have no rotational freedom as they are part of the peptide bond, and the residues that they belong to are further constrained by hydrogen bonds involving their peptide CO groups. Thus, the Ser114 CO hydrogen bonds to Arg119 NH_2, the Gly115 CO hydrogen bonds to Ser118 NH and Arg119 NE, the Leu116 CO hydrogen bonds to Asn10 ND2, and the Trp117 CO hydrogen bonds to the NH group of Tyr100. As many of the groups on the protein that are hydrogen bond donors to the sulfate oxygens have no rotational freedom and belong to residues that are constrained in the protein, the precise fold of the polypeptide backbone is critical for the geometry of the Cys-MR ligand binding site.

Although the sulfate portion of 4-SO_4-GalNAc is negatively charged, the binding site on Cys-MR does not have a positive electrostatic potential (Fig. 2 E) and the sulfate group does not interact with positively charged residues or counter ions (Table II). These features are shared with another protein that binds sulfate, the sulfate-binding protein (SBP) from *Salmonella typhimurium* (43). Based on studies of SBP and other proteins that bind charged ligands without using counter ions, a mechanism for stabilizing isolated charges on groups sequestered in proteins was proposed (44). In this mechanism, hydrogen bonds, especially those formed with highly polarized peptide units, stabilize charges on buried groups, which are further dispersed by hydrogen bond arrays. The sulfate binding interactions of both Cys-MR and SBP conform to three common features of proteins that bury isolated negative charges: (1) the interactions are mediated primarily through hydrogen bonds rather than salt bridges or counter ions; (2) most of the hydrogen bonds involve peptide NH groups, and (3) the peptide NH groups associated with the sulfate group are coupled to a hydrogen bond array within the protein.

The Cys-MR ligand binding site differs in many respects from the binding sites of other proteins with β-trefoil folds, including the ricin B chain and FGF, both of which bind carbohydrates. The galactose binding site of the ricin B chain is in lobe I instead of lobe III (Fig. 1, A and C). However, in common with the interaction between the galactose ring of 4-SO_4-GalNAc and Cys-MR, the galactose ligand of ricin B is stacked against a tryptophan residue (Trp37), although, compared with the Cys-MR tryptophan, the opposite face of the ricin tryptophan ring contacts the galactose (Fig. 2 D). While FGF binds a sulfated carbohydrate using a region of its structure analogous to the Cys-MR ligand binding site (the strand 11–12 region of lobe III; Fig. 1, A and C), the FGF site is a fairly flat, positively charged surface that uses electrostatic interactions for ligand recognition (45), by contrast with the neutral pocket that constitutes the Cys-MR binding site, in which electrostatic interactions do not play a role in ligand binding.

Carbohydrate Binding Specificity of Cys-MR. The structure of Cys-MR bound to 4-SO_4-GalNAc can be used to rationalize the observed binding specificity of Cys-MR for sulfated carbohydrates. Biochemical experiments (19) concluded that Cys-MR binds to GalNAc or galactose when it is sulfated on the C3 or C4 position, but only weakly if the sulfate is at the C6 position (19). The lack of discrimination between sulfated galactose and sulfated *N*-acetylgalactosamine is explained by the absence of contacts between the N-acetyl group of 4-SO_4-GalNAc and Cys-MR (Table II). To understand the mechanism that the position of the sulfate is discriminated, we used models of sulfated GalNAc molecules (Fig. 3 A) to predict the interactions between Cys-MR and GalNAc that is sulfated at the C3 or C6 positions and compared these with the interaction observed in the Cys-MR/4-SO_4-GalNAc complex structure. For each carbohydrate model, we positioned the sulfate group in the Cys-MR binding site in the position observed in the crystal structure of Cys-MR complexed with 4-SO_4-GalNAc and allowed rotational freedom about the galactose C6-sulfate oxygen and the galactose C5–C6 bonds. In the case of 3-SO_4-GalNAc, the galactose ring can be fit into the binding site without steric hindrance, making extensive van der Waals contacts with Trp117. The galactose ring of 6-SO_4-GalNAc also fits into the binding site without steric hindrance, but cannot be rotated into a position in which the sugar ring makes extensive van der Waals contacts with Trp117.

The same binding study concluded that Cys-MR recognizes sulfated GalNAc when it is in a terminal, rather than an internal, position within an oligosaccharide chain (19). For example, adding a terminal uronic acid with a β1-3 linkage to a 4-SO_4-GalNAc group prevents binding to Cys-MR. From the mode of binding revealed by the crystal structure, this result is expected since linking a saccharide group to the C3 position of 4-SO_4-GalNAc should prevent access of the sulfate group to the binding site (Fig. 2 C). By contrast, linkage of a sugar to the C1 position of sulfated galactose should not generally affect binding to Cys-MR. Thus, Cys-MR binds chondroitin sulfates that have a terminal 4-sulfated GalNAc with a β1-4 linkage to iduronic acid or to glucuronic acid and also to sulfated blood group carbohydrates with a terminal 3-sulfated galactose joined with a β1-4 or β1-3 linkage to *N*-acetylglucosamine (19). As the hydroxyl group attached to C1 is exposed to solvent in the Cys-MR/4-SO_4-GalNAc structure and is predicted to be exposed on 3-SO_4-GalNAc bound to Cys-MR (Fig. 3 A), it should be possible to link another sugar ring at this position without creating steric hindrance. Thus, it is predicted that the sugar adjoining the terminal sulfated sugar would not contact Cys-MR unless it contained bulky substitutions, thus rationalizing the observation that the identity of the sugar linked to the terminal sulfated carbohydrate did not generally affect the ability to bind to Cys-MR (19).

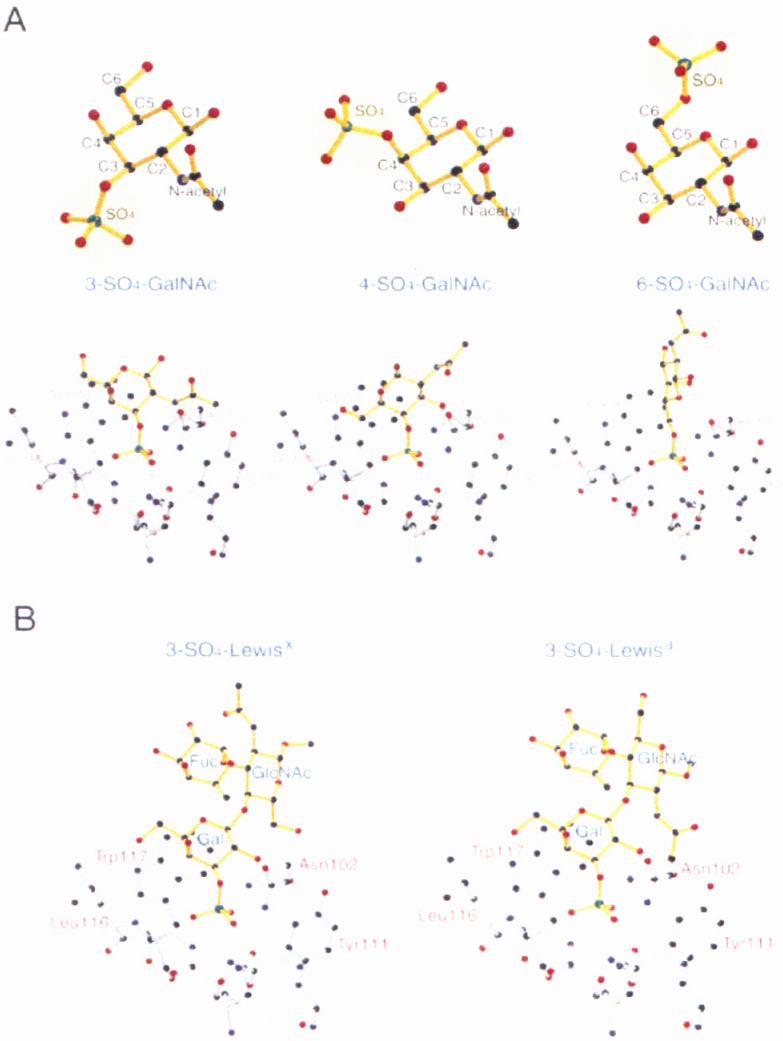


Figure 3. Models of sulfated carbohydrates binding to Cys-MR. (A) Modeled interactions between sulfated GalNAc molecules and Cys-MR. The predicted interactions of Cys-MR with 3-SO_4-GalNAc and 6-SO_4-GalNAc were compared with the observed structure of Cys-MR bound to 4-SO_4-GalNAc. Top: A sulfate group was attached to positions 3 or 6 of the galactose ring of a model of GalNAc (derived from the structure of Cys-MR complexed with 4-SO_4-GalNAc) to produce the sulfated carbohydrate structures shown. Bottom: For each model, the sulfate group was anchored in the Cys-MR binding site as observed in the native and 4-SO_4-GalNAc–complexed Cys-MR structures, and the GalNAc portion was rotated about the galactose C6-sulfate oxygen and C5–C6 bonds to manually arrive at the best fit. (B) Modeled interactions between sulfated Lewis^x and Lewis^a molecules and Cys-MR. The structure of 3-SO_4-Lewis^x was derived from the structures of Lewis^x alone (reference 46) and sulfated Lewis^x bound to a selectin-like mutant of mannose-binding protein (reference 47). The structure of 3-SO_4-Lewis^a was modeled from the 3-SO_4-Lewis^x structure as described (reference 48). 3-SO_4-Lewis^x and 3-SO_4-Lewis^a were modeled into the Cys-MR binding site with their 3-SO_4-galactose rings positioned as in the model of 3-SO_4-GalNAc bound to Cys-MR (A). In this orientation of 3-SO_4-Lewis^a bound to Cys-MR, the *N*-acetyl group on the galactose ring is predicted to clash with Asn102.

Other ligands of Cys-MR include sulfated oligosaccharides of the blood group Lewis^x and Lewis^a series. Lewis^x and Lewis^a saccharides that include a sulfate group at the C3 position of the terminal galactose ring (3-SO_4-Lewis^x and 3-SO_4-Lewis^a) bind Cys-MR, with 3-SO_4-Lewis^x binding slightly (approximately three-fold) better than 3-SO_4-Lewis^a (19). To model the interaction between Cys-MR and sulfated Lewis^x, we used a model of 3-SO_4-Lewis^x derived from the structures of Lewis^x alone (46) and sulfated Lewis^x bound to a selectin-like mutant of mannose-binding protein (47). The structure of 3-SO_4-Lewis^a was modeled from the 3-SO_4-Lewis^x structure as described (48). The 3-SO_4-galactose portions of 3-SO_4-Lewis^x and 3-SO_4-Lewis^a (Fig. 3 B) were superimposed with the model of 3-SO_4-GalNAc bound to Cys-MR (Fig. 3 A), allowing rotational freedom about the sulfo–galactose bond. In the case of 3-SO_4-Lewis^x, a stacking interaction between Trp117, the sulfated galactose ring, and the fucose ring is predicted (Fig. 3 B). When 3-SO_4-Lewis^a is positioned similarly in the binding site, the *N*-acetyl portion of the glucose ring of Lewis^a clashes with Cys-MR residue Asn102 (Fig. 3 B), which may account for the observed preference for 3-SO_4-Lewis^x over 3-SO_4-Lewis^a (19).

Discussion

MR is a carbohydrate-binding receptor that contains regions with distinct carbohydrate-recognition properties.

The first involves binding of C-type lectin CRDs to glycans that terminate with mannose, fucose, or *N*-acetylglucosamine residues. The second involves binding of the NH_2-terminal Cys-MR domain to acidic glycans such as those terminating with 4-sulfated *N*-acetylgalactosamine or 3-sulfated galactose. Ligands for the CRDs include carbohydrates on the surfaces of infectious agents; thus, MR is thought to serve as a pattern-recognition receptor in the innate immune response and to deliver carbohydrate-containing antigens to MHC molecules for antigen presentation in the adaptive immune response (4). In addition, the MR CRDs have recently been shown to bind to pituitary hormones, most likely through interactions involving terminal mannose or core fucose residues (49). Carbohydrate-containing ligands for the Cys-MR domain include endogenous proteins such as sulfated pituitary hormones (16–18) and molecules expressed on cells in the spleen and germinal centers (19, 22). The dual carbohydrate specificity of MR gives it the potential to serve as a bridge between ligands containing nonacidic glycans such as mannose and those containing acidic sulfated glycans. For example, macrophages secrete a soluble form of MR (50), which has been proposed to direct mannose-bearing antigens to cells expressing Cys-MR ligand(s) at sites where humoral immune responses take place (22, 50).

The two types of carbohydrate-binding regions of MR have different structures and sugar-binding properties. The eight CRDs of MR are predicted to fold into structures resembling C-type lectin domains (4, 14), as recently confirmed by the crystal structure of MR CRD4.[2] The CRD fold contains α-helical and β-strand secondary structural elements and regions with no ordered secondary structure that include the calcium-binding loops (14, 42). Calcium is required for the integrity of the fold and for binding carbohydrates, such that a calcium ion forms direct coordination bonds with mannose, fucose, *N*-acetylglucosamine, and glucose molecules. Protein–ligand interactions in CRDs include hydrogen bonds between the sugar hydroxyls and protein side chains (14, 42). The fold of the NH_2-terminal Cys-MR region, the other carbohydrate-binding portion of MR, bears no resemblance to CRDs. Instead, Cys-MR is a member of the β-trefoil family of proteins, which are approximately three-fold symmetric structures composed of 12 β strands (34). Other β-trefoil proteins, such as acidic FGF (aFGF) and the trefoil region of the ricin B chain, bind carbohydrates, but their binding sites and mechanisms of carbohydrate recognition differ from those of Cys-MR (Fig. 1 A). Thus, the Cys-MR binding site and residues critical for ligand recognition could not have been predicted from structural comparisons, but were revealed in the crystal structure of Cys-MR complexed with 4-SO_4-GalNAc. The binding site is a neutral pocket that accommodates the sulfate group, with numerous hydrogen bonds between the sulfate oxygens and protein atoms, primarily main chain NH groups, which serve to disperse the negative charge. Although the galactose ring of the ligand makes specific contacts with Cys-MR, the majority of interactions involve the sulfate group, rationalizing the specificity of Cys-MR for acidic carbohydrate ligands such as sulfated galactose (4).

Identification of the residues in the Cys-MR binding site facilitates comparisons with the sequences of cysteine-rich regions of other proteins in the MR family. The cysteine-rich domains of other members of the MR family, including DEC-205 (10) and the phospholipase A2 receptor (7–9), contain a deletion in their sequences in the region analogous to that used by Cys-MR for binding 4-SO_4-GalNAc (Fig. 1 C). Thus, these receptors are not predicted to bind sulfated carbohydrates using in a manner analogous to Cys-MR, consistent with the observation that this region of DEC-205 does not bind lutropin (19). Together with the structure of CRD4,[2] the structure of Cys-MR provides a structural framework to evaluate sugar recognition by the two types of carbohydrate-binding portions of the MR family of proteins.

We thank S. Gordon and L. Martínez-Pomares for the Cys-MR-Fc expression vector, G. Hathaway and the California Institute of Technology's Protein/Peptide Micro Analytical Laboratory for peptide analyses, M.J. Bennett, A.P. Yeh, L.M. Sanchez, H.J. Chiu, S. Ding, and M. Williamson for crystallographic assistance, and W.I. Weis, M.E. Taylor, and members of the Bjorkman laboratory for critical reading of the manuscript. Coordinates have been submitted to the Protein Data Bank (available at www.rcsb.org/pdb/ under accession nos. IDQG and IDQO).

This work was supported by the Howard Hughes Medical Institute (P.J. Bjorkman and M.C. Nussenzweig) and grants from the National Institutes of Health (M.C. Nussenzweig). Y. Liu was a recipient of Ferguson Fund predoctoral Fellowship and Teaching Assistantship (California Institute of Technology).

Submitted: 7 December 1999
Revised: 27 January 2000
Accepted: 4 February 2000

References

1. Janeway, C.A., Jr. 1992. The immune system evolved to distinguish infectious nonself from non-infectious self. *Immunol. Today.* 13:11–16.
2. Takahashi, K., M.J. Donovan, R.A. Rogers, and R.A.B. Ezekowitz. 1998. Distribution of murine mannose receptor expression from early embryogenesis through to adulthood. *Cell Tissue Res.* 292:311–323.
3. Linehan, S.A., L. Martínez-Pomares, P.D. Stahl, and S. Gordon. 1999. Mannose receptor and its putative ligands in normal murine lymphoid and nonlymphoid organs: in situ expression of mannose receptor by selected macrophages, endothelial cell, perivascular microglia, and mesangial cells, but not dendritic cells. *J. Exp. Med.* 189:1961–1972.
4. Stahl, P.D., and R.A.B. Ezekowitz. 1998. The mannose receptor is a pattern recognition receptor involved in host defense. *Curr. Opin. Immunol.* 10:50–55.
5. Sallusto, F., M.I. Cella, C. Danieli, and A. Lanzavecchia. 1995. Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. *J. Exp. Med.* 182:389–400.
6. Prigozy, T.I., P.A. Sieling, D. Clemens, P.L. Stewart, S.M.

Behar, S.A. Porcelli, M.B. Brenner, R.L. Modlin, and M. Kronenberg. 1997. The mannose receptor delivers lipoglycan antigens to endosomes for presentation to T cells by CD1b molecules. *Immunity*. 6:187–197.

7. Ishizaki, J., K. Hanasaki, K. Higashino, J. Kishimo, N. Kikuchi, O. Ohara, and H. Arita. 1994. Molecular cloning of pancreatic group I phospholipase A2 receptor. *J. Biol. Chem.* 269:12156–12162.
8. Lambeau, G., P. Ancian, J. Barhanin, and M. Lazdunski. 1994. Cloning and expression of a membrane receptor for secretory phospholipases A_2. *J. Biol. Chem.* 269:1575–1578.
9. Ancian, P., G. Lambeau, M.-G. Madei, and M. Lazdunski. 1995. The human 180 kDa receptor for secretory phospholipases A_2. *J. Biol. Chem.* 270:8963–8970.
10. Jiang, W.P., W.J. Swiggard, C. Heufler, M. Peng, A. Mirza, R.M. Steinman, and M.C. Nussenzweig. 1995. The receptor DEC-205 expressed by dendritic cells and thymic epithelial cells is involved in antigen processing. *Nature*. 375:151–155.
11. Wu, K., J. Yuan, and L.A. Lasky. 1996. Characterization of a novel member of the macrophage mannose receptor type C lectin family. *J. Biol. Chem.* 271:21323–21330.
12. Taylor, M.E., K. Bezouska, and K. Drickamer. 1992. Contribution to ligand binding by multiple carbohydrate recognition domains in the macrophage mannose receptor. *J. Biol. Chem.* 267:1719–1726.
13. Taylor, M.E., and K. Drickamer. 1993. Structural requirements for high affinity binding of complex ligands by the mannose receptor. *J. Biol. Chem.* 268:399–404.
14. Weis, W.I., M.E. Taylor, and K. Drickamer. 1998. The C-type lectin superfamily in the immune system. *Immunol. Rev.* 163:19–34.
15. Mullin, N.P., P.G. Hitchen, and M.E. Taylor. 1997. Mechanism of Ca_2^+ and monosaccharide binding to a C-type carbohydrate-recognition domain of the macrophage mannose receptor. *J. Biol. Chem.* 272:5668–5681.
16. Fiete, D., and J.U. Baenziger. 1997. Isolation of the SO_4-4-GalNAcβ1, 4GlcNAcβ1, 2Manα-specific receptor from rat liver. *J. Biol. Chem.* 272:14629–14637.
17. Fiete, D., M.C. Beranek, and J.U. Baenziger. 1997. The macrophage endothelial cell mannose receptor cDNA encodes a protein that binds oligosaccharides terminating with SO_4-4-GalNAcβ1,4GlcNAcβ or Man at independent sites. *Proc. Natl. Acad. Sci. USA*. 94:11256–11261.
18. Fiete, D.J., M.C. Beranek, and J.U. Baenziger. 1998. A cysteine-rich domain of the "mannose" receptor mediates GalNAc-4-SO_4 binding. *Proc. Natl. Acad. Sci. USA*. 95:2089–2093.
19. Leteux, C., W. Chai, R.W. Loveless, C.-T. Yuen, L. Uhlin-Hansen, Y. Combarnous, M. Jankovic, S.C. Maric, Z. Misulovin, M.C. Nussenzweig, and T. Feizi. 2000. The cysteine-rich domain of the macrophage mannose receptor is a multispecific lectin which recognizes chondroitin sulfates A and B and sulfated oligosaccharides of blood-group Lewis^a and Lewis^x types in addition to the sulfated *N*-glycans of luteinizing hormone. *J. Exp. Med.* 191:1117–1126.
20. Baenziger, J.U., S. Kumar, R.M. Brodbeck, P.L. Smith, and M.C. Beranek. 1992. Circulatory half-life but not interaction with the lutropin chorionic-gonadotropin receptor is modulated by sulfation of bovine lutropin oligosaccharides. *Proc. Natl. Acad. Sci. USA*. 89:334–338.
21. Smith, P.L., G.R. Bousfield, S. Kuman, D. Fiete, and J.U. Baenziger. 1993. Equine lutropin and chorionic gonadotropin bear oligosaccharides terminating with SO_4-4-GalNAc and Sia-α-2,3Gal, respectively. *J. Biol. Chem.* 268:795–802.
22. Martínez-Pomares, L., M. Kosco-Vilbois, E. Darley, P. Tree, S. Herren, J.-Y. Bonnefoy, and S. Gordon. 1996. Fc chimeric protein containing the cysteine-rich domain of the murine mannose receptor binds to macrophages from splenic marginal zone and lymph node subcapsular sinus and to germinal centers. *J. Exp. Med.* 184:1927–1937.
23. Taylor, M.E., J.T. Conary, M.R. Lennartz, P.D. Stahl, and K. Drickmer. 1990. Primary structure of the mannose receptor contains multiple motifs resembling carbohydrate-recognition domains. *J. Biol. Chem.* 265:12156–12162.
24. Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307–326.
25. De La Fortelle, E., and G. Bricogne. 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* 276:472–494.
26. Abrahams, J.P., and A.G.W. Leslie. 1996. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D.* 52:30–42.
27. Kleywegt, G.J., and T.A. Jones. 1997. Template convolution to enhance or detect structural features in macromolecular electron density maps. *Acta Crystallogr. D.* 53:179–185.
28. Jones, T.A., and M. Kjeldgaard. 1997. Electron density map interpretation. *Methods Enzymol.* 277:173–208.
29. Brünger, A.T., P.D. Adams, G.M. Clore, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, et al. 1998. Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D.* 54:905–921.
30. Brünger, A.T. 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*. 355: 472–475.
31. Kraulis, P.J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.
32. Merritt, E.A., and M.E.P. Murphy. 1994. Raster3D version 2.0, a program for photorealistic molecular graphics. *Acta Crystallogr. D.* 50:869-873.
33. Nicholls, A., R. Bharadwaj, and B. Honig. 1993. GRASP: graphical representation and analysis of surface properties. *Biophys. J.* 64:A166.
34. Murzin, A.G., A.M. Lesk, and C. Chothia. 1992. Beta-trefoil fold patterns of structure and sequence in the Kunitz inhibitors, interleukins-1β and 1α and fibroblast growth factors. *J. Mol. Biol.* 223:531–543.
35. McLachlan, A.D. 1979. Three-fold structural pattern in the soybean trypsin inhibitory (Kunitz). *J. Mol. Biol.* 133:557–563.
36. Rutenber, E., and J.D. Robertus. 1991. Structure of ricin B-chain at 2.5 Å resolution. *Proteins*. 10:260–269.
37. Zhu, X., H. Komiya, A. Chirino, S. Faham, G.M. Fox, T. Arakawa, B.T. Hsu, and D.C. Rees. 1991. Three-dimensional structures of acidic and basic fibroblast growth factors. *Science*. 251:90–93.
38. Finzel, B.C., L.L. Clancy, D.R. Holland, S.W. Muchmore, K.D. Watenpaugh, and H.M. Einspahr. 1989. Crystal structure of recombinant human interleukin-1β at 2.0 Å resolution. *J. Mol. Biol.* 209:779–791.
39. Graves, B.J., M.H. Hatada, W.A. Hendrickson, J.K. Miller, V.S. Madison, and Y. Satow. 1990. Structure of interleukin-1α at 2.7 Å resolution. *Biochemistry*. 29:2679–2684.

40. Habazetti, J., D. Gondol, R. Wiltscheck, J. Otlewski, M. Schleicher, and T.A. Holak. 1992. Structure of hisactophilin is similar to interleukin-1β and fibroblast growth factor. *Nature*. 359:855–858.
41. Harris, N., L.L. Peters, E.M. Eicher, M. Rits, D. Raspberry, Q.G. Eichbaum, M. Super, and R.A. Ezekowitz. 1994. The exon-intron structure and chromosomal localization of the mouse macrophage mannose receptor gene Mrc1: identification of a ricin-like domain at the N-terminus of the receptor. *Biochem. Biophys. Res. Commun.* 198:682–692.
42. Weis, W.I., and K. Drickamer. 1996. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* 65:441–473.
43. Pflugrath, J.W., and F.A. Quiocho. 1988. The 2 Å resolution structure of the sulfate-binding protein involved in active transport in *Salmonella typhimurium*. *J. Mol. Biol.* 200:163–180.
44. Quiocho, F.A., J.S. Sack, and N.K. Vyas. 1987. Stabilization of charges on isolated ionic groups sequestered in proteins by polarized peptide units. *Nature*. 329:561–564.
45. DiGabriele, A.D., I. Lax, D.I. Chen, C.M. Svahn, M. Jaye, J. Schlessinger, and W.A. Hendrickson. 1998. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature*. 393:812–817.
46. Perez, S., N. Mouhous-Riou, N.E. Nifant'ev, Y.E. Tsvetkov, B. Bachet, and A. Imberty. 1996. Crystal and molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis^x trisaccharide. *Glycobiology*. 6:537–542.
47. Ng, K.K.-S., and W.I. Weis. 1997. Structure of a selectin-like mutant of mannose binding protein complexed with sialylated and sulfated Lewis^x oligosaccharides. *Biochemistry*. 36: 979–988.
48. Kogelberg, H., T.A. Frenkiel, S.W. Homans, A. Lubineau, and T. Feizi. 1996. Conformational studies on the selectin and natural killer cell receptor ligands sulfo- and sialyl-lacto-*N*-fucopentaoses (SuLNFPII and SLNFPII) using NMR spectroscopy and molecular dynamics simulations. Comparisons with the nonacidic parent molecule LNFPII. *Biochemistry*. 35:1954–1964.
49. Simpson, D.Z., P.G. Hitchen, E.L. Elmhirst, and M.E. Taylor. 1999. Multiple interactions between pituitary hormones and the mannose receptor. *Biochem. J.* 343:403–411.
50. Martínez-Pomares, L., J.A. Mahoney, R. Káposzta, S.A. Linehan, P.D. Stahl, and S. Gordon. 1998. A functional soluble form of the murine mannose receptor is produced by macrophages in vitro and is present in mouse serum. *J. Biol. Chem.* 273:23376–23380.
51. Hodel, A., S.-H. Kim, and A.T. Brünger. 1992. Model bias in macromolecular crystal structures. *Acta Crystallogr. A.* 48: 851–858.
52. Kolatkar, A.R., and W.I. Weis. 1996. Structural basis of galactose recognition by C-type animal lectins. *J. Biol. Chem.* 271:6679–6685.

Chapter 3

The Molecular Mechanism of Sulfated Carbohydrate Recognition by the Cysteine-rich Domain of Mannose Receptor

Running title: Cys-MR/sulfated carbohydrate binding

The Molecular Mechanism of Sulfated Carbohydrate Recognition by the Cysteine-rich Domain of Mannose Receptor

Yang Liu[*], Ziva Misulovin[§] and Pamela J. Bjorkman[*‡¶]

[*]Division of Biology 156-29, [‡]Howard Hughes Medical Institute, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA.

[§]Molecular Immunology, Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021-6399, USA.

[¶]To whom reprint requests and page proofs should be addressed:

Pamela J. Bjorkman

Division of Biology 156-29

Caltech

Pasadena, CA 91125

Phone: 626 395-8350; Fax: 626 792-3683; E-mail: bjorkman@cco.caltech.edu

Summary

The mannose receptor (MR) binds foreign and host ligands through interactions with their carbohydrates. Two portions of MR have distinct carbohydrate recognition properties. One is conferred by the amino-terminal cysteine-rich domain (Cys-MR), which plays a critical role in binding sulfated glycoproteins including pituitary hormones. The other is achieved by tandemly arranged C-type lectin domains that facilitate carbohydrate-dependent uptake of infectious microorganisms. This dual carbohydrate binding specificity enables MR to bind ligands by interacting with both sulfated and non-sulfated polysaccharide chains. We previously determined crystal structures of Cys-MR complexed with 4-SO_4-GalNAc and with an unidentified ligand resembling Hepes (N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]). In continued efforts to elucidate the mechanism of sulfated carbohydrate recognition by Cys-MR, we characterized the binding affinities between Cys-MR and potential carbohydrate ligands using a fluorescence-based assay. We find that Cys-MR binds sulfated carbohydrates with relatively high affinities (K_D = 0.1 mM to 1.0 mM) compared to the affinities of other lectins. Cys-MR also binds Hepes with a K_D of 3.9 mM, consistent with the suggestion that the ligand in the original Cys-MR crystal structure is Hepes. We also determined crystal structures of Cys-MR complexed with 3-SO_4-Lewis^x, 3-SO_4-Lewis^a, and 6-SO_4-GalNAc at 1.9 Å, 2.2 Å, and 2.5 Å resolution, respectively, and the 2.0 Å structure of Cys-MR that had been treated to remove Hepes. The conformation of the Cys-MR binding site is virtually identical in all Cys-MR crystal structures, suggesting that Cys-MR does not undergo conformational changes upon ligand binding. The structures are used to rationalize the binding affinities derived from the biochemical studies and to elucidate the molecular mechanism of sulfated carbohydrate recognition by Cys-MR.

Key words: β-trefoil protein, fluorescence spectrophotometery, mannose receptor, pituitary hormones, sulfated carbohydrates.

Introduction

The macrophage and epithelial cell mannose receptor (MR) is a 180 kDa type I transmembrane protein that contains the following extracellular regions: a membrane-distal amino-terminal cysteine-rich domain (Cys-MR), a fibronectin type II-like domain, and eight tandem C-type lectin carbohydrate-recognition domains (CRDs) (Stahl and Ezekowitz, 1998). Cys-MR and the tandem CRD regions have been shown to have distinct carbohydrate binding properties, with Cys-MR binding to sulfated carbohydrates (Fiete et al., 1998; Leteux et al., 2000) and the CRD regions recognizing terminal mannose, fucose, and *N*-acetylglucosamine (GlcNAc) (Taylor et al., 1992; Taylor and Drickamer, 1993). The extracellular domain of MR is connected to a transmembrane domain and a small cytoplasmic domain, which contains an endocytic signal motif that is critical for MR mediated endocytosis and phagocytosis (Nicolas et al., 1995; Stahl et al., 1980). MR is the prototype member of a family of multilectin receptors whose members share similar structural characteristics. Other members of the family include the phospholipase A2 receptor (Ancian et al., 1995; Ishizaki et al., 1994; Lambeau et al., 1994), the dendritic cell receptor DEC-205 (Jiang et al., 1995), and an MR-like receptor (Wu et al., 1996).

MR is widely expressed on macrophages (Ezekowitz et al., 1990; Taylor et al., 1990), immature dendritic cells (Sallusto et al., 1995), and a variety of endothelial and epithelial cells (Stahl and Ezekowitz, 1998). It has diverse immune-related functions, including roles in innate and adaptive immunity. In the innate immune response, MR facilitates macrophage phagocytosis of microorganisms, such as bacteria, yeast and parasites, via recognition of the polysaccharides on their cell walls (Stahl and Ezekowitz, 1998). As part of the adaptive immune response, MR helps to direct macrophages or MR-bearing cells toward germinal centers, and a soluble form of MR transports mannose-bearing antigens to the sites where B cell maturation occurs (Martinez-Pomares et al., 1996; Martinez-Pomares et al., 1998). In addition, MR is thought to participate in the antigen processing and presentation pathways of

MHC class II molecules (Sallusto et al., 1995) and the class I MHC-like protein CD1b (Prigozy et al., 1997).

In addition to its immune functions, MR on hepatic endothelial cells and Kupffer cells is thought to play a critical role in controlling the circulatory half-life of pituitary hormones such as luteinizing hormone (LH) and thyroid-stimulating hormone (TSH) (Fiete et al., 1997; Fiete et al., 1991) by rapidly removing these hormones from the circulation, thus preventing desensitization of the LH receptor on target cells (Baenziger et al., 1992). Abundant expression of MR on hepatic endothelial cells and Kupffer cells facilitates the hormonal uptake and removal processes (Fiete et al., 1991). MR binds pituitary hormones in a pH dependent manner at pH values $\geq$ 5.5 (Fiete et al., 1991; Simpson et al., 1999). Both the Cys-MR and CRD regions of MR interact with pituitary hormones, with Cys-MR binding to sulfated carbohydrate chains on the hormone (Fiete et al., 1991) and the CRDs interacting with non-sulfated carbohydrates (Simpson et al., 1999).

Biochemical studies have been used to define the specificity of sulfated carbohydrate recognition by Cys-MR. Ligands of Cys-MR include: 1) N-glycans terminating with 4-sulfated-*N*-acetylgalactosamine (4-SO_4-GalNAc) on pituitary hormones (Fiete et al., 1998); 2) chondroitin 4-sulfate chains terminating with 4-SO_4-GalNAc on lymphocytes, monocytes and some other hematopoietic cells; 3) and sulfated blood group chains terminating with 3-sulfated-galactose (3-SO_4-Gal), such as 3-SO_4-Lewis^x and 3-SO_4-Lewis^a (Leteux et al., 2000). The crystal structures of mouse Cys-MR complexed with 4-SO_4-GalNAc and an unknown ligand revealed an approximately three-fold symmetric β-trefoil shape in which the sulfate groups of the 4-SO_4-GalNAc and unknown ligands were accommodated in a neutral pocket within the third lobe (Liu et al., 2000) (Figure 1A). To further characterize the molecular mechanism of sulfated carbohydrate recognition by Cys-MR, we determined binding affinities between Cys-MR and carbohydrate ligands using a fluorescence-based assay, yielding K_D values between 0.1 and 1 mM. We also determined the crystal structures of unliganded Cys-MR at 2.0 Å resolution and Cys-MR structures complexed with 3-SO_4-Lewis^x, 3-SO_4-Lewis^a,

or 6-SO_4-GalNAc to resolutions of 1.9 Å, 2.2 Å, and 2.5 Å, respectively. The relatively high affinity of Cys-MR for sulfated carbohydrates and the binding preferences for different sulfated ligands are rationalized using the crystallographic results.

Results

Equilibrium binding studies

The crystal structure of the Cys-MR/4-SO_4-GalNAc complex revealed hydrophobic stacking interactions between the galactose ring of the ligand and Cys-MR Trp117 (Liu et al., 2000) (Figure 1B), suggesting that changes in intrinsic tryptophan fluorescence could be used to monitor ligand binding. Upon excitation at 280 nm, tyrosines and tryptophans emit fluorescence with maxima at 303 nm and 348 nm, respectively (Teale and Weber, 1957). Changes in the environment of tryptophans can be evaluated by monitoring fluorescence emission at 345 nm. By determining the degree of fluorescence change as a function of ligand concentration, we determined equilibrium dissociation constants (K_Ds) for the binding of five different sulfated carbohydrates (4-SO_4-GalNAc, 3-SO_4-Lewis^x, 3-SO_4-Lewis^a, 6-SO_4-GalNAc, and 6-SO_4-Gal) to Cys-MR (Figure 2) (see Methods). Binding constants ranged from 0.11 mM for 4-SO_4-GalNAc to 1.0 mM for 6-SO_4-Gal. By contrast, we see no significant change in fluorescence intensity at 345 nm, and hence no detectable binding, when Cys-MR is incubated with non-sulfated carbohydrates such as GalNAc or 6-PO_4-Gal. Other studies comparing binding of Cys-MR to immobilized pentasaccharide versions of 3-SO_4-Lewis^x and 3-SO_4-Lewis^a found weaker binding of 3-SO_4-Lewis^a compared with 3-SO_4-Lewis^x (Leteux et al., 2000). The similar K_Ds derived in this study for binding of 3-SO_4-Lewis^a and 3-SO_4-Lewis^a (Figure 2) may reflect the greater rotational freedom of the trisaccharide Lewis acids in solution compared with the immobilized glycoconjugates used previously.

We previously suggested that the unknown ligand found in the crystal structure of native Cys-MR could represent a molecule of Hepes (N-[2-Hydroxyethyl]piperazine-N'-[2-

ethanesulfonic acid]), which was present at a concentration of 10 mM in the crystallization buffer. We find that Hepes binds to Cys-MR more weakly than the sulfated carbohydrates (K_D = 3.9 mM), but with a high enough affinity to support the suggestion that it could occupy the Cys-MR binding site in the absence of other ligands.

pH dependent binding of Cys-MR to 4-SO_4-GalNAc

MR is an endocytic receptor that mediates ligand binding at the cell surface and delivers bound ligands to acidified endosomes (Prigozy et al., 1997; Stahl et al., 1980). MR binding to mannose/GlcNAc/fucose-terminating glyco-conjugates and to SO_4-GalNAcβ1,4GlcNAcβ1,2Manα-derivatized bovine serum albumin is pH dependent (Fiete et al., 1991; Lennartz et al., 1987), such that MR binds at pH values between 5.5 and 7.5 and releases at pH 5. In addition, MR and a truncated form of MR lacking Cys-MR and the fibronectin type II repeat bind LH and TSH with pH dependence, with maximal binding at pH 6-7 and a sharp decrease below pH 5.5 (Simpson et al., 1999). In these studies, the role of Cys-MR in pH dependent ligand binding was not directly evaluated, although the studies with truncated MR demonstrate that the CRDs alone show pH dependent binding. In order to address if the Cys-MR region also binds ligands with pH dependence, we determined K_Ds for the binding of 4-SO_4-GalNAc to Cys-MR as a function of pH. We find that binding affinity of 4-SO_4-GalNAc to Cys-MR is fairly constant between pH 4.5 and 7.5 and does not drop significantly until the pH is lowered to 4 (Figure 3). This result suggests that the Cys-MR region of MR does not play a significant role in the release of pituitary hormones or other ligands in endosomal compartments, since these compartments range in pH from 5.0 to 6.0 (Stryer, 1995).

Crystal structure of "ligand-free" Cys-MR

The Cys-MR ligand binding site was previously described based upon a 1.7 Å crystal structure of Cys-MR, which contained an unidentified ligand that co-crystallized with the

protein (*Liu et al., 2000*), and a 2.2 Å structure of Cys-MR bound to 4-SO_4-GalNAc (Liu et al., 2000). We now identify the unknown ligand as Hepes, a component in the crystallization buffer, based upon the fit between the molecular structure of Hepes and the electron density of the ligand (Figure 4A) and the demonstration that Hepes binds to Cys-MR with a K_D of 3.9 mM (Figure 2).

Cys-MR folds into a globular shape composed of 12 antiparallel β-strands arranged in a pattern with approximate three-fold internal symmetry. The binding site for both 4-SO_4-GalNAc and Hepes is located in a neutral pocket formed on one side by residues 111-116 in the loop between strands 11 and 12, and on the other side by sidechains of residues including Asn102, Trp117 and Asn99. The sulfate portion of 4-SO_4-GalNAc and the SO_3 group of Hepes point into the pocket, making extensive hydrogen bonds with backbone atoms within the strand 11-12 loop. The SO_4 or SO_3 group of each ligand makes the majority of interactions with Cys-MR, although the galactose ring of 4-SO_4-GalNAc stacks against the Trp117, and hydroxyls on the galactose ring hydrogen bond with Asn102 and Leu116. No significant conformational changes were observed between the two structures, but since both contained a ligand, we could not address whether unliganded Cys-MR changes upon binding.

In the present study, we attempted to remove Hepes from Cys-MR in order to solve the structure of unliganded Cys-MR. We soaked Cys-MR crystals in mother liquor without Hepes or ammonium sulfate for 7-10 days and determined the crystal structure to 2.0 Å (Table 1). The refined model of the Cys-MR structure derived from the soaked crystals is virtually identical to the original 1.7 Å Cys-MR structure (root mean square (rms) deviation of 0.14 Å for all Cα atoms), except that there is clear density at the Cys-MR binding site for only the SO_3 group of Hepes in maps contoured at 3 σ (Figure 4B). By comparison, in the original structure, the SO_3 group and part of the piperazine ring of Hepes were apparent in electron density maps contoured at 3 σ (Figure 4A). In the new structure, the SO_3 group has a refined occupancy of 70% compared with 100% for the SO_3 group in the original Cys-MR/Hepes structure. This result indicates that we were only partially successful in removing Hepes from

the Cys-MR binding site. Further soaking of Cys-MR crystals in Hepes-free mother liquor destroys the crystals (Y.L. and P.J.B., data not shown).

Despite our inability to completely remove Hepes from the Cys-MR crystals, we can use the new crystal structure to address the potential for conformational change in Cys-MR upon ligand binding. Since the occupancy of Hepes in the soaked crystals is not 100%, some of the Cys-MR molecules are empty. The structure derived from the soaked crystals does not show disorder or high temperature factors in the vicinity of the binding site, as would be expected if the empty Cys-MR molecules had a different structure than the ligand-occupied Cys-MR structure. Thus we conclude that ligand binding is unlikely to induced conformational changes at the Cys-MR binding site or in the rest of the Cys-MR structure.

Structures of Cys-MR complexed with sulfated Lewis acids

Sulfated forms of Lewis^x and Lewis^a were previously shown to bind to Cys-MR (Leteux et al., 2000) and a model of their binding to Cys-MR was constructed using information from the Cys-MR/4-SO_4-GalNAc structure (Liu et al., 2000). Here we present crystal structures of 3-SO_4-Lewis^x and 3-SO_4-Lewis^a bound to Cys-MR at 1.9 Å and 2.2 Å resolution, respectively (Table 1; Figure 4C, D). The protein portions of the structures are nearly identical to the 1.7 Å Cys-MR/Hepes structure (rms deviations of 0.25 Å and 0.24 Å, respectively, for Cα atom comparisons with Cys-MR/3-SO_4-Lewis^x and Cys-MR/3-SO_4-Lewis^a). The sidechain and mainchain atoms of residues at the ligand binding sites in the new structures are unchanged from the positions observed in the other Cys-MR structures, including Cys-MR/Hepes, Cys-MR/4-SO_4-GalNAc, and the soaked Cys-MR crystals.

The positions of the bound ligands were determined unambiguously using electron density derived from Fo-Fc simulated annealing omit maps (Figure 4C, D), although the average temperature factor for 3-SO_4-Lewis^a (B_{avg} = 74.6 Å^2) is higher than that for Lewis^x (B_{avg} = 54.6 Å^2). In maps contoured at 1 σ (Figure 4C, D), most portions of the ligands are well-defined, and in maps contoured at 2 σ, the 3-SO_4-Gal and the fucose ring of 3-SO_4-Lewis^x

and 3-SO_4-*Gal and* part of the fucose of 3-SO_4-Lewis^a are visible (data not shown). Using electron density maps contoured at 1 σ, we were able to place the sulfated galactose, *N*-acetyl glucosamine, and fucose rings of both ligands with certainty. The 3-SO_4-Gal portion of each ligand is positioned similarly in the Cys-MR binding site, with the SO_4 group fitting into the same position as the SO_4 group of 4-SO_4-GalNAc in the Cys-MR/4-SO_4-GalNAc structure (Liu et al., 2000) (Figure 5). Thus both the Lewis^x and Lewis^a sulfate groups form six hydrogen bonds with Cys-MR, including two involving sidechain atoms of Asn99 and Asn102, and four involving the main chain NH groups of Ser114, Gly115, Leu116 and Trp117 (Figure 5). Consistent with the earlier structural studies (Liu et al., 2000), the extensive hydrogen bonding network between Cys-MR and the sulfate group plays a critical role in anchoring sulfated carbohydrates by dispersing the negative charge of sulfate group throughout the neutral Cys-MR ligand binding pocket.

The galactose rings of the 3-SO_4-Gal portion of sulfated Lewis^x and Lewis^a use their non-polar faces to stack against the sidechain of Trp117, as previously observed for the galactose ring of 4-SO_4-GalNAc (Liu et al., 2000). However, the galactose ring of 4-SO_4-GalNAc makes more interactions with Trp117 than the galactose rings of the 3-SO_4-Gal part of Lewis^x and Lewis^a. Whereas 4-SO_4-GalNAc makes 14 van der Waals interactions with Trp117 using carbon atoms C3, C4, C5, and C6 of its galactose ring (Liu et al., 2000), the galactose rings of the 3-SO_4-Gal portion of Lewis^x and Lewis^a forms only seven van der Waals interactions using carbon atoms C1, C3, and C5. Thus, 4-SO_4-GalNAc has stronger ring stacking interactions with Cys-MR than the 3-SO_4-Gal portions of 3-SO_4-Lewis^x and 3-SO_4-Lewis^a. This result is consistent with the observation that 4-SO_4-GalNAc has a slightly higher binding affinity (0.11 mM) for Cys-MR than 3-SO_4-Lewis^x and 3-SO_4-Lewis^a (0.2 mM and 0.19 mM, respectively).

Using the results of NMR studies suggesting that Lewis^a and Lewis^x have similar structures (Kogelberg et al., 1996) and the structure of 3-SO_4-Lewis^x when bound to a mannose binding protein mutant (Ng and Weis, 1997), we can address whether the Lewis^x and Lewis^a

structures undergo conformational changes in response to binding Cys-MR. We find that the structures of 3-SO_4-Lewis^x bound to a mannose-binding protein mutant and bound to Cys-MR are similar (0.81 Å rms deviation for all atoms; 0.24 Å rms deviation for the carbon and oxygen atoms in sugar rings). The other ligand, 3-SO_4-Lewis^a, may undergo conformational changes upon binding to Cys-MR, since the 3-SO_4-Lewis^a molecule in the Cys-MR/3-SO_4-Lewis^a complex structure shows significant differences from the 3-SO_4-Lewis^x structure, such that its GlcNAc and fucose portions are bent away from Asn102 (Figure 4F). This structural change is required to avoid a clash between the N-acetyl portion of the Lewis^a glucose ring and Asn102 of Cys-MR, as predicted by modeling of the sulfated Lewis^a in the Cys-MR structure (Liu et al., 2000).

Structure of Cys-MR complexed with 6-SO_4-GalNAc

Previous biochemical studies demonstrated that Cys-MR binds GalNAc or galactose that is sulfated on the C3 or C4 position, but only weak binding is observed when the sulfate group is attached to the C6 position (Leteux et al., 2000). Here we demonstrate that the binding affinity of 6-SO_4-GalNAc for Cys-MR is ~9-fold weaker than the affinity between 4-SO_4-GalNAc and Cys-MR (Figure 2). In our previous modeling studies, we noted that 6-SO_4-GalNAc can be positioned into the Cys-MR binding site with its sulfate group in the position of the 4-SO_4-GalNAc sulfate group, but that the galactose ring of 6-SO_4-GalNAc cannot be rotated into a position in which the sugar ring makes extensive van der Waals contacts with Trp 117 (Liu et al., 2000). In order to address this prediction, we soaked Cys-MR crystals in 6-SO_4-GalNAc and solved the crystal structure to 2.5 Å resolution. Once again, the protein portion of the structure shows no significant changes compared with the Cys-MR/Hepes structure (0.39 Å rms deviation for Cα atoms). However, the only density at the ligand binding site corresponds to a sulfate group. There is no electron density for the GalNAc portion of the ligand in Fo-Fc difference maps contoured at 2σ (Figure 4E). Since 6-SO_4-GalNAc has a higher binding affinity for Cys-MR than Hepes (K_D = 0.94 mM compared with

3.9 mM), it should bind to Cys-MR under the conditions of the soaking experiment (10 mM 6-SO_4-GalNAc). We therefore assume that 6-SO_4-GalNAc is bound to Cys-MR, but that the GalNAc portion of 6-SO_4-GalNAc is disordered, consistent with the prediction that the galactose ring cannot stack against Trp 117 (Liu et al., 2000).

Discussion

MR is a multilectin receptor that is divided into domains with different carbohydrate recognition properties. The N-terminal cysteine-rich domain (Cys-MR) binds N-linked glycans on pituitary hormones terminating with 4-SO_4-GalNAc and sulfated glycans on cells in the spleen and germinal centers (Leteux et al., 2000; Martinez-Pomares et al., 1996), whereas the tandemly arranged C-type lectin CRDs bind glycans that terminate with mannose, fucose, or GlcNAc (Taylor et al., 1992; Taylor and Drickamer, 1993). This dual carbohydrate binding specificity allows MR to achieve high-affinity binding to LH and TSH by simultaneously recognizing terminal 4-SO_4-GalNAc and mannose residues (Simpson et al., 1999). Crystal structures of Cys-MR (Liu et al., 2000) and the principal mannose-binding domain of MR (CRD4) (Feinberg et al., 2000) reveal different tertiary structures and modes of ligand binding. Cys-MR has a β-trefoil fold, consisting of twelve antiparallel β-strands arranged with approximate three fold internal symmetry, and binds sulfated carbohydrates in a neutral binding pocket within the third lobe (Liu et al., 2000). CRD4 has a canonical C-type lectin fold, which includes β-sheet and α-helical secondary structures and a calcium binding site that is required for carbohydrate recognition (Feinberg et al., 2000).

We previously reported the crystal structure of Cys-MR bound to 4-SO_4-GalNAc and an unidentified ligand resembling a molecule of Hepes (Liu et al., 2000), and here present continued biochemical and structural studies aimed at defining the molecular mechanism for sulfated carbohydrate recognition by Cys-MR. Using a fluorescence-based assay, we determined binding affinities between Cys-MR and various carbohydrates. We characterized the binding of Cys-MR to three types of carbohydrates: 1) sulfated carbohydrates containing

an SO_4 group linked to the C3 or C4 position of a sugar ring, such as 4-SO_4-GalNAc, 3-SO_4-Lewis^x, and 3-SO_4-Lewis^a, 2) sulfated carbohydrates containing an SO_4 group linked to the C6 position of a sugar ring, such as 6-SO_4-GalNAc and 6-SO_4-Gal, and 3) non-sulfated carbohydrates, such as GalNAc and 6-PO_4-Gal. Those in the first group bind to Cys-MR with the highest binding affinities (K_D = 0.11 mM for 4-SO_4-GalNAc and ~0.2 mM for 3-SO_4-Lewis^x and 3-SO_4-Lewis^a) (Figure 2). Crystal structures of Cys-MR complexed with 4-SO_4-GalNAc, 3-SO_4-Lewis^x, and 3-SO_4-Lewis^a reveal that the sulfate and galactose portions of each of the compounds are positioned similarly in the Cys-MR binding site, such that the sulfate group forms extensive hydrogen bonds to main- and sidechain atoms of Cys-MR and the galactose ring stacks against Trp117 (Figure 5). Compounds in the second carbohydrate group show lower binding affinities for Cys-MR (0.94 mM and 1.0 mM for 6-SO_4-GalNAc and 6-SO_4-Gal, respectively). The crystal structure of Cys-MR/6-SO_4-GalNAc shows electron density only for the sulfate portion of the ligand (Figure 4E), consistent with the prediction that the galactose ring cannot rotate into a position in which it can stack against Trp117, and therefore binds more weakly to Cys-MR. Neither of the compounds in the third group of carbohydrates showed detectable binding to Cys-MR, confirming the requirement for a sulfate group and revealing that the slightly larger phosphate group is not accommodated within the Cys-MR binding site. We also used the fluorescence-based binding assay to demonstrate that Cys-MR binds to Hepes (K_D = 3.9 mM), providing additional evidence that the unidentified ligand in the original Cys-MR crystal structure is a molecule of Hepes, which was a component in the crystallization buffer (Liu et al., 2000). Hepes, which contains an SO_3 group linked to a piperazine ring, is positioned in the Cys-MR binding site such that the SO_3 group makes all of the same hydrogen bonds as the SO_4 groups of 4-SO_4-GalNAc or the sulfated Lewis acids, but the piperazine ring cannot stack optimally against Trp117 and is not completely visible in electron density maps (Figure 4A) (Liu et al., 2000). Attempts to soak Hepes out of the crystals were only partially successful, resulting in crystals with Hepes

bound at less than 100% occupancy (Figure 4B), but are consistent with the suggestion that Cys-MR does not undergo conformational changes upon ligand binding.

MR is highly expressed on hepatic endothelial cells (~600,000 receptors/cell) and is believed to play a crucial role in uptake of circulating LH (Fiete et al., 1991). Both the Cys-MR and CRDs regions of MR are involved in recognizing LH (Simpson et al., 1999). The ligand binding affinity of Cys-MR (~0.1 mM) is higher than the affinity of a single CRD for its ligand, which are generally in the mM range (e.g., the K_D for CRD4 binding to α-methyl mannoside is 2.4 mM (Mullin et al., 1997)). The weak affinities of individual CRDs for their ligands can be compensated for by avidity effects created through the use of tandemly arranged domains. Thus the affinity of the CRD4-8 portion of MR for high mannose-containing structures ranges from 20 to 100 nM (Taylor et al., 1992). The higher ligand binding affinity of Cys-MR compared to individual CRDs is important for ligand recognition, since there is only one Cys-MR domain in MR and MR family members (Stahl and Ezekowitz, 1998). The overall binding affinity of MR to LH (K_D = 140 nM) reflects the effects of multiple interactions between MR and sulfated and non-sulfated polysaccharides on LH (Fiete et al., 1991; Simpson et al., 1999). The crystal structures and binding studies presented here contribute to an overall understanding of the mechanism of sulfated carbohydrate recognition by the Cys-MR region of MR, and facilitate evaluation of the functions of MR in pituitary hormone regulation and immunity.

Materials and Methods

Fluorescence-based binding assay

The Cys-MR domain of mouse MR was expressed and purified as described (Liu et al., 2000). 4-SO_4-GalNAc, 6-SO_4-Gal, 6-SO_4-GalNAc, 3-SO_4-Lewis^x, GalNAc and 6-PO_4-Gal were purchased from Sigma Chemical Co. 3-SO_4-Lewis^a was purchased from Glyko.com. We used a Hitachi F-4500 fluorescence spectrophotometer to record fluorescence spectra of Cys-MR in the presence and absence of carbohydrate at room temperature. Fluorescence spectra ranging from 300 nm to 400 nm were obtained after excitation at 280 ± 5nm. The emission spectrum of Cys-MR alone showed a peak at 345 nm, thus fluorometric titrations of Cys-MR with carbohydrates were evaluated using the change of fluorescence intensity at 345nm. For fluorometric titrations with 4-SO_4-GalNAc, 6-SO_4-Gal, 6-SO_4-GalNAc, Hepes, GalNAc and 6-PO_4-Gal, 200 µl of 4.0 µM Cys-MR in 20 mM Tris pH 7.4, 150 mM NaCl was incubated with each carbohydrate at the following concentrations: 0 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.4 mM, 0.6 mM, 0.8 mM, 1.0 mM, 2.0 mM, 3.0 mM, 4.0 mM, 5.0 mM, and 6.0 mM. For titrations of Cys-MR with 3-SO_4-Lewis^x and 3-SO_4-Lewis^a (which were available in smaller quantities), concentrated carbohydrate solutions were added to 300 µl of 4.0 µM Cys-MR to produce final carbohydrate concentrations of 0 mM, 0.08 mM, 0.16 mM, 0.32 mM, 0.64 mM, 1.28 mM, 2.56 mM, and 5.12 mM and the fluorescence intensity measurements were adjusted to account for changes in protein concentration.

The dissociation constant (K_D) for the binding of Cys-MR (C) to a carbohydrate ligand (L) is expressed as:

$$K_D = [C]\,[L] / [C{\bullet}L] = ([C]_0 - [C{\bullet}L])([L]_0 - [C{\bullet}L]) / [C{\bullet}L]$$

$$[C{\bullet}L] = (K_D + [L]_0 + [C]_0 - ((K_D + [L]_0 + [C]_0)^2 - 4[L]_0[C]_0)^{1/2}) / 2$$

where ($[C]_0$ and $[L]_0$ are the total concentrations of Cys-MR and carbohydrate, respectively, [C], [L], and [C•L] are the equilibrium concentrations of Cys-MR, carbohydrate ligand, and Cys-MR complexed with carbohydrate, respectively.

Assuming that the change in fluorescence intensity is directly proportional to the concentration of the Cys-MR/carbohydrate complex, the observed fluorescence intensity (F_{obs}) is expressed as:

$$F_{obs} = F_0 - \Delta F = F_0 - \Delta f_{C\bullet L}[C\bullet L]$$

$$= F_0 - \Delta f_{C\bullet L}(K_D + [L]_0 + [C]_0 - ((K_D + [L]_0 + [C]_0)^2 - 4[L]_0[C]_0)^{1/2}) / 2$$

where F_0 is the initial fluorescence intensity of Cys-MR in the absence of ligand, and ΔF is the total fluorescence intensity change for the Cys-MR/carbohydrate complex.

The fluorescence intensity change ($|F_{obs} - F_0|$) is represented as a percentage of the initial fluorescence intensity, F_0, of Cys-MR at 345 nm:

$$|F_0 - F_{obs}|/F_0 = \Delta f_{C\bullet L}(K_D + [L]_0 + [C]_0 - ((K_D + [L]_0 + [C]_0)^2 - 4[L]_0[C]_0)^{1/2}) / (2F_0)$$

K_D and $\Delta f_{C\bullet L}$ values were determined assuming a 1:1 binding interaction for $[C]_0 = 4.0$ μM by non-linear regression analysis of the values of $|F_0 - F_{obs}|/F_0$ for a series of ligand concentrations ($[L]_0$) using the program KaleidaGraph. Using the derived K_D and $\Delta f_{C\bullet L}$ values, the maximum value of $|F_0 - F_{obs}|/F_0$ for each carbohydrate ligand was calculated and represented as $|F_0 - F_{max}|/F_0$. The percent ligand bound at each concentration was calculated as $|F_0 - F_{obs}|/|F_0 - F_{max}|$.

pH dependence of Cys-MR binding to 4-SO_4-GalNAc

Cys-MR was exchanged into 20 mM of the following buffers for determinations of K_D values as a function of pH: sodium formate (pH 3.0, 3.5), sodium acetate (pH 4.0, 4.5), cacodylic acid (pH 5.0, 5.5, 6.0, 6.5), and Tris-HCl (pH 7.0, 7.5). All buffers included 150 mM NaCl. Concentrated 4-SO_4-GalNAc was added to 300 μl of 4.0 μM Cys-MR to produce final carbohydrate concentrations of 0 mM, 0.08 mM, 0.16 mM, 0.32 mM, 0.64 mM, 1.28 mM, 2.56 mM, and 5.12 mM, and changes in fluorescence intensity were adjusted to account for changes in protein concentration.

Crystallization and data collection

Cys-MR crystals (space group $P2_12_12_1$; one molecule per asymmetric unit) were grown by hanging drop vapor diffusion at 4°C. Cys-MR (40 mg/ml in 10 mM Hepes pH 7.4, 10 mM NaCl) was mixed with an equal volume of reservoir solution containing 0.2 M $(NH_4)_2SO_4$ and 30% (w/v) PEG8000. To remove Hepes from the Cys-MR binding site, crystals were washed in 35% (w/v) PEG 8000, then soaked in 35% PEG 8000 and suspended over a reservoir of 32% (w/v) PEG 8000, 0.2M $(NH_4)_2SO_4$. To obtain complex crystals, Cys-MR crystals were incubated in equilibrated hanging drops containing 0.2M $(NH4)_2SO_4$, 32% (w/v) PEG 8000, and 10 mM 3-SO_4-Lewis^x, 3-SO_4-Lewis^a, or 6-SO_4-GalNAc. After 7-14 days at 4°C, crystals were transferred to a cryoprotectant solution (40% PEG 8000 for unliganded crystals; 32% PEG 8000, 0.2M $(NH4)_2SO_4$, 10% glycerol for complex crystals) and were flash cooled in liquid nitrogen. All data sets were collected at –150°C from single crystals using an R-AXIS IV detector (Molecular Structures Corporation) mounted on a Rigaku R200 rotating anode generator. Data sets were processed and scaled with the HKL package (Otwinowski and Minor, 1997).

Structure determination and refinement

The structures of unliganded Cys-MR and Cys-MR complexed with carbohydrate ligands were solved by molecular replacement using the 1.7 Å Cys-MR/Hepes structure (Liu et al., 2000) (PDB code 1DQG) as a search model. All models were refined with CNS (Brünger et al., 1998) using bulk solvent correction, the maximum likelihood target function, refinement of individual temperature (B) factors, and protocols to minimize R_{free} (Brünger, 1992) as described (Liu et al., 2000). Water molecules were built into peaks >3σ in difference Fourier maps. Statistics for the refinement and model geometry are summarized in Table 1.

Acknowledgments

We thank Ten Feizi, Christine Leteux, Heide Kojelberg, and members of the Bjorkman laboratory for critical reading of the manuscript, and Michel Nussenzweig for contributions to earlier phases of the project. Coordinates will be deposited in the Protein Data Bank.

Figure 1. *Structure* of Cys-MR bound to 4-SO_4-GalNAc

(A) Ribbon diagram of the Cys-MR structure with lobes I, II, and III indicated in different colors. Disulfide bonds are yellow and 4-SO_4-GalNAc is shown in ball-and-stick representation. (B) Space filling representation of Cys-MR (calculated using GRASP; (Nicholls et al., 1993)) showing location of the two tryptophans in the structure. Trp117 is buried upon ligand binding.

Figure 2. Fluorescence-based binding assay

The percent of ligand bound is plotted versus the log of the concentration of added ligand. Data points shown correspond to the average of triplicate measurements at each concentration and bars denote standard deviations. Non-linear regression analysis was used to fit the data to a 1:1 binding model and to derive values for the K_D and the maximum binding response. Best fit curves to triplicate data points are shown in solid colors.

Figure 3. pH dependence of Cys-MR binding to 4-SO_4-GalNAc

K_Ds for Cys-MR binding to 4-SO_4-GalNAc were determined for the pH range from 3.0 to 7.5.

Figure 4. Close-up of Cys-MR ligand binding site derived from Cys-MR/ligand crystal structures

Panels A-E. Cys-MR residues are shown in ball-and-stick representation (gray) with the ligand (yellow) superimposed on cyan electron density derived from Fo-Fc annealed omit maps (Hodel et al., 1992). The resolution of each structure is indicated in parentheses. (A) Cys-MR/Hepes structure (Liu et al., 2000). Contour level for electron density: 3 σ. (B) Cys-MR structure derived after soaking crystals to remove Hepes. Contour level for electron density: 3 σ. (C) Cys-MR/3-SO_4-Lewis^x structure. Contour level for electron density: 1.0 σ. (D) Cys-MR/3-SO_4-Lewis^a structure. Contour level for electron density: 1 σ. (E) Cys-MR/6-

SO_4-GalNAc *structure*. Only the SO_4 group of 6-SO_4-GalNAc is visible in electron density maps. Contour level for electron density: 2 σ. (F) Superposition of Cys-MR binding site residues and ligands from the Cys-MR/3-SO_4-Lewis^x (green) and Cys-MR/3-SO_4-Lewis^a (purple).

Figure 5. Comparison of Cys-MR binding to 3-SO_4-Gal and 4-SO_4-GalNAc

Stereo view of the interactions between Cys-MR and the 3-SO_4-Gal portion of 3-SO_4-Lewis^x and 4-SO_4-GalNAc. Hydrogen bonds between ligand and protein atoms are indicated by dotted green lines.

Table 1. Data collection, phasing, and refinement statistics

Data Set	Resolution (Å)	Number of Observations	Unique Reflections	% Complete*	R_{merge} (%)‡	I/σI	Unit Cell a (Å), b (Å), c (Å)
"ligand-free"	25-2.0 (2.07-2.0)	87031	11434	98.3 (85.8)	3.6 (15.7)	29.5 (6.7)	39.23, 40.94, 97.88
Cys-MR/ 3-SO_4-Lewis^x	25-1.9 (1.96-1.9)	108934	13916	95.4 (86.6)	3.6 (12.3)	28.2 (9.0)	39.59, 41.30, 100.40
Cys-MR/ 3-SO_4-Lewis^a	25-2.2 (2.28-2.2)	64935	8862	96.5 (89.1)	3.4 (8.5)	31.5 (12.4)	39.61, 41.16, 99.65
Cys-MR/ 6-SO_4-GalNAc	25-2.5 (2.59-2.5)	59661	6176	96.7 (95.1)	3.7 (9.1)	29.9 (11.7)	39.60, 41.28, 100.16

Refinement

	"ligand-free"	Cys-MR/ 3-SO_4-Lewis^x	Cys-MR/ 3-SO_4-Lewis^a	Cys-MR/ 6-SO_4-GalNAc
Resolution (Å)	25.0-2.0	25.0-1.9	25.0-2.2	25.0-2.5
Reflections in working set |F|>0	10443	12433	7965	5515
Reflections in test set |F|>0	676	793	505	349
R_{cryst}/R_{free} (%)¶	19.9/22.1	21.3/23.7	20.8/23.0	21.9/24.7
rmsd bond length (Å)	0.0055	0.008	0.0067	0.0067
rmsd angles (deg)	1.37	1.40	1.27	1.36
Number of atoms				
protein	1085	1085	1085	1085
water	168	130	127	56
ligand	4	40	40	5
Ramachandran plot				
most favored region (%)	84.4	86.9	83.6	85.0
additional allowed region (%)	15.6	12.3	16.4	14.2
generously allowed region (%)	0.0	0.8	0.0	0.8
disallowed region (%)	0.0	0.0	0.0	0.0

Values in parentheses indicate the high-resolution shells. *Complete is the number of independent reflections/total theoretical number. ‡R_{merge} (I)= (Σ|I(i) - <I(**h**)>|/ΣI(i)), where I(i) is the *i*th observation of the intensity of the hkl reflection and <I> is the mean intensity from multiple measurements of the h,k,l reflection. ¶R_{cryst} (F) = $\sum_h||\mathbf{F_{obs}(h)}| - |\mathbf{F_{calc}(h)}||/\sum_h|\mathbf{F_{obs}(h)}|$, where $|\mathbf{F_{obs}(h)}|$ and $|\mathbf{F_{calc}(h)}|$ are the observed and calculated structure factor amplitudes for the h,k,l reflection. R_{free} is calculated over reflections in a test set not included in atomic refinement.

Abbreviations

CRD, carbohydrate recognition domain; Cys-MR, the cysteine-rich domain of mannose receptor; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; Hepes, N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid); K_D, equilibrium dissociation constant; LH, luteinizing hormone; MHC, major histocompatibility complex; MR, mannose receptor; TSH, thyroid-stimulating hormone; rms, root mean square.

References:

Ancian, P., Lambeau, G., Mattei, M. G., and Lazdunski, M. (1995). The human 180-kDa receptor for secretory phospholipases A2. Molecular cloning, identification of a secreted soluble form, expression, and chromosomal localization. J. Biol. Chem. *270*, 8963-8970.

Baenziger, J. U., Kumar, S., Brodbeck, R. M., Smith, P. L., and Beranek, M. C. (1992). Circulatory half-life but not interaction with the lutropin/chorionic gonadotropin receptor is modulated by sulfation of bovine lutropin oligosaccharides. Proc. Natl. Acad. Sci. USA *89*, 334-338.

Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905-921.

Ezekowitz, R. A., Sastry, K., Bailly, P., and Warner, A. (1990). Molecular characterization of the human macrophage mannose receptor: demonstration of multiple carbohydrate recognition-like domains and phagocytosis of yeasts in Cos-1 cells. J. Exp. Med. *172*, 1785-1794.

Feinberg, H., Torgerson, D., Drickamer, K., and Weis, W. I. (2000). Mechanism of pH-dependent N-acetylgalactosamine binding by a functional mimic of the hepatocyte asialoglycoprotein receptor. J. Biol. Chem. *275*, 21539-21548.

Fiete, D., Beranek, M. C., and Baenziger, J. U. (1997). The macrophage/endothelial cell mannose receptor cDNA encodes a protein that binds oligosaccharides terminating with SO4-4-GalNAcbeta1,4GlcNAcbeta or Man at independent sites. Proc. Natl. Acad. Sci. USA *94*, 11254-11261.

Fiete, D., Srivastava, V., Hindsgau, O., and Baenziger, J. U. (1991). A hepatic reticuloendothelial cell receptor specific for SO4-4GalNAc beta 1,4GlcNAc beta 1,2Man alpha that mediates rapid clearance of lutropin. Cell *67*, 1103-1110.

Fiete, D. J., Beranek, M. C., and Baenziger, J. U. (1998). A cysteine-rich domain of the "mannose" receptor mediates GalNAc-4-SO4 binding. Proc. Natl. Acad. Sci. USA. *95*, 2089-2093.

Hodel, A., Kim, S.-H., and Brünger, A. T. (1992). Model bias in macromolecular crystal structures. Acta crystallogr. *A48*, 851-858.

Ishizaki, J., Hanasaki, K., Higashino, K., Kishino, J., Kikuchi, N., Ohara, O., and Arita, H. (1994). Molecular cloning of pancreatic group I phospholipase A2 receptor. J. Biol. Chem. *269*, 5897-5904.

Jiang, W., Swiggard, W. J., Heufler, C., Peng, M., Mirza, A., Steinman, R. M., and Nussenzweig, M. C. (1995). The receptor DEC-205 expressed by dendritic cells and thymic epithelial cells is involved in antigen processing. Nature *375*, 151-155.

Kogelberg, H., Frenkiel, T. A., Homans, S. W., Lubineau, A., and Feizi, T. (1996). Conformational studies on the selectin and natural killer cell receptor ligands sulfo- and sialyl-lacto-N-fucopentaoses (SuLNFPII and SLNFPII) using NMR spectroscopy and molecular

dynamics simulations. Comparisons with the nonacidic parent molecule LNFPII. Biochemistry *35*, 1954-1964.

Lambeau, G., Ancian, P., Barhanin, J., and Lazdunski, M. (1994). Cloning and expression of a membrane receptor for secretory phospholipases A2. J. Biol. Chem. *269*, 1575-1578.

Lennartz, M. R., Cole, F. S., Shepherd, V. L., Wileman, T. E., and Stahl, P. D. (1987). Isolation and characterization of a mannose-specific endocytosis receptor from human placenta. J. Biol. Chem. *262*, 9942-9944.

Leteux, C., Chai, W., Loveless, R. W., Yuen, C. T., Uhlin-Hansen, L., Combarnous, Y., Jankovic, M., Maric, S. C., Misulovin, Z., Nussenzweig, M. C., and Feizi, T. (2000). The cysteine-rich domain of the macrophage mannose receptor is a multispecific lectin that recognizes chondroitin sulfates A and B and sulfated oligosaccharides of blood group Lewis(a) and Lewis(x) types in addition to the sulfated N-glycans of lutropin. J. Exp. Med. *191*, 1117-1126.

Liu, Y., Chirino, A. J., Misulovin, Z., Leteux, C., Feizi, T., Nussenzweig, M. C., and Bjorkman, P. J. (2000). Crystal structure of the cysteine-rich domain of mannose receptor complexed with a sulfated carbohydrate ligand. J. Exp. Med. *191*, 1105-1116.

Martinez-Pomares, L., Kosco-Vilbois, M., Darley, E., Tree, P., Herren, S., Bonnefoy, J. Y., and Gordon, S. (1996). Fc chimeric protein containing the cysteine-rich domain of the murine mannose receptor binds to macrophages from splenic marginal zone and lymph node subcapsular sinus and to germinal centers. J. Exp. Med. *184*, 1927-1937.

Martinez-Pomares, L., Mahoney, J. A., Kaposzta, R., Linehan, S. A., Stahl, P. D., and Gordon, S. (1998). A functional soluble form of the murine mannose receptor is produced by macrophages in vitro and is present in mouse serum. J. Biol. Chem. *273*, 23376-23380.

Mullin, N. P., Hitchen, P. G., and Taylor, M. E. (1997). Mechanism of Ca2+ and monosaccharide binding to a C-type carbohydrate-recognition domain of the macrophage mannose receptor. J Biol Chem. *272*, 5668-5681.

Ng, K., and Weis, W. (1997). Structure of a selectin-like mutant of mannose-binding protein complexed with sialylated and sulfated Lewis(x) oligosaccharides. Biochemistry *36*, 979-988.

Nicholls, A., Bharadwaj, R., and Honig, B. (1993). GRASP - Graphical representation and analysis of surface properties. Biophys. J. *64*, A166-A166.

Nicolas, J. P., Lambeau, G., and Lazdunski, M. (1995). Identification of the binding domain for secretory phospholipases A2 on their M-type 180-kDa membrane receptor. J. Biol. Chem. *270*, 28869-28873.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. *276*, 307-326.

Prigozy, T. I., Sieling, P. A., Clemens, D., Stewart, P. L., Behar, S. M., Porcelli, S. A., Brenner, M. B., Modlin, R. L., and Kronenberg, M. (1997). The mannose receptor delivers lipoglycan antigens to endosomes for presentation to T cells by CD1b molecules. Immunity *6*, 187-197.

Sallusto, F., Cella, M., Danieli, C., and Lanzavecchia, A. (1995). Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. J. Exp. Med. *182*, 389-400.

Simpson, D. Z., Hitchen, P. G., Elmhirst, E. L., and Taylor, M. E. (1999). Multiple interactions between pituitary hormones and the mannose receptor. Biochem. J. *343*, 403-411.

Stahl, P., Schlesinger, P. H., Sigardson, E., Rodman, J. S., and Lee, Y. C. (1980). Receptor-mediated pinocytosis of mannose glycoconjugates by macrophages: characterization and evidence for receptor recycling. Cell *19*, 207-215.

Stahl, P. D., and Ezekowitz, R. A. (1998). The mannose receptor is a pattern recognition receptor involved in host defense. Curr. Opin. Immunol. *10*, 50-55.

Stryer, L. (1995). Biochemistry, IV Edition (New York: Freeman).

Taylor, M. E., Bezouska, K., and Drickamer, K. (1992). Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. J. Biol. Chem. *267*, 1719-1726.

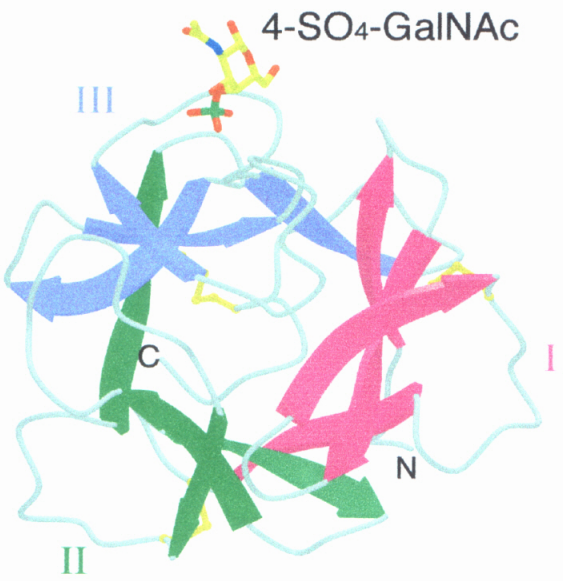
Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. D., and Drickamer, K. (1990). Primary structure of the mannose receptor contains multiple motifs resembling carbohydrate-recognition domains. J. Biol. Chem. *265*, 12156-12162.

Taylor, M. E., and Drickamer, K. (1993). Structural requirements for high affinity binding of complex ligands by the macrophage mannose receptor. J. Biol. Chem. *268*, 399-404.

Teale, F. J. W., and Weber, G. (1957). Ultraviolet fluorescence of the aromatic amino acids. Biochem. J. *65*, 476-482.

Wu, K., Yuan, J., and Lasky, L. A. (1996). Characterization of a novel member of the macrophage mannose receptor type C lectin family. J. Biol. Chem. *271*, 21323-21330.

A



B

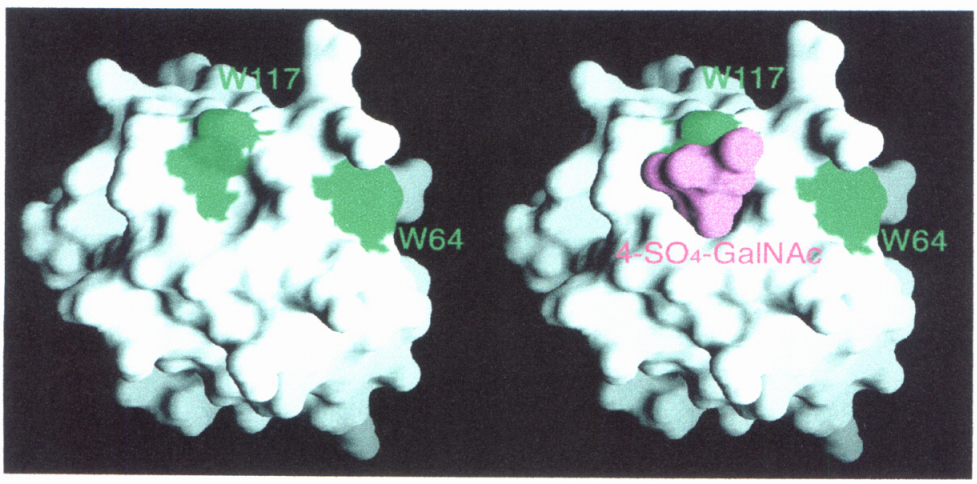


Figure 1. Liu et al.

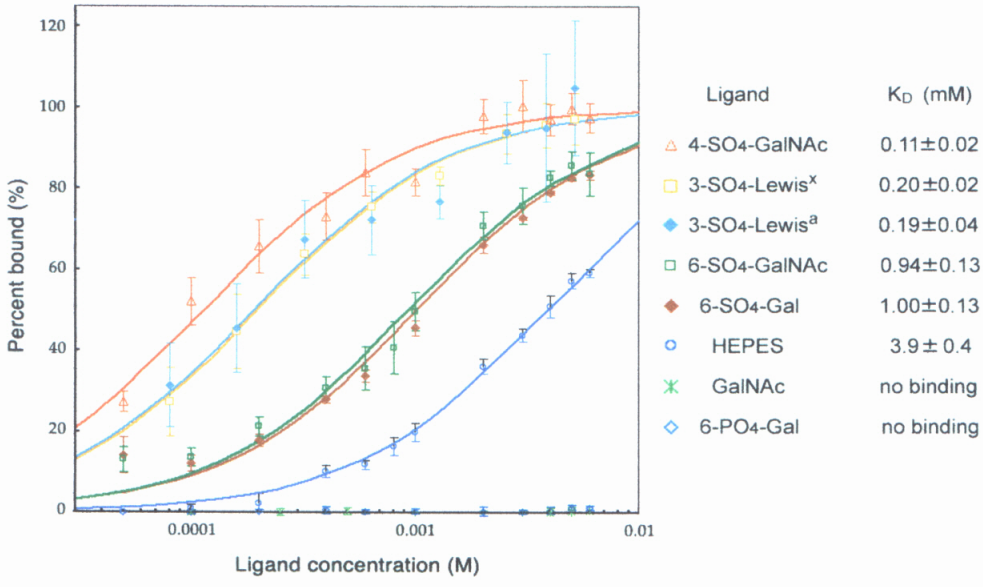


Figure 2. Liu et al.

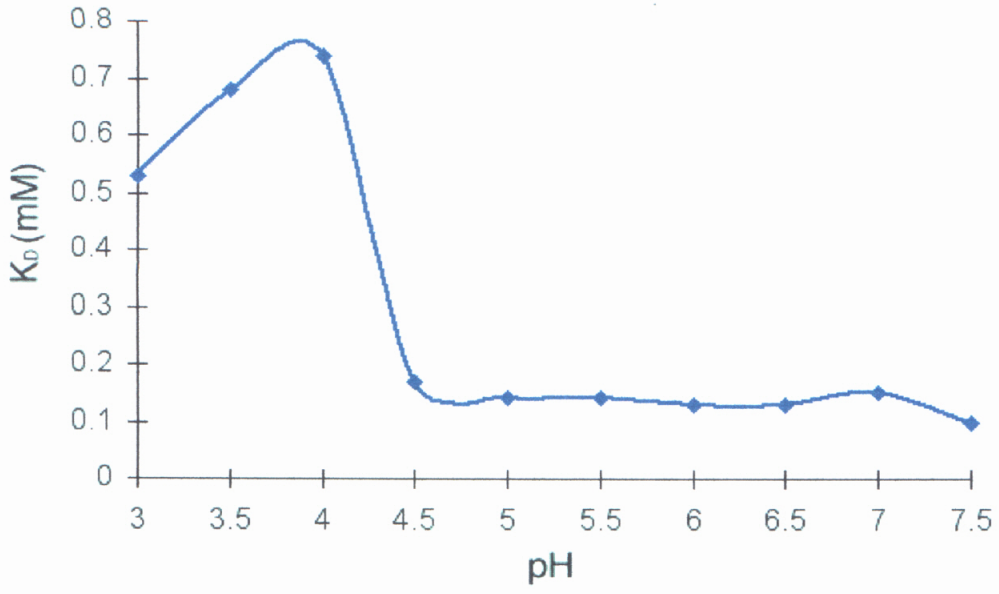


Figure 3. Liu et al.

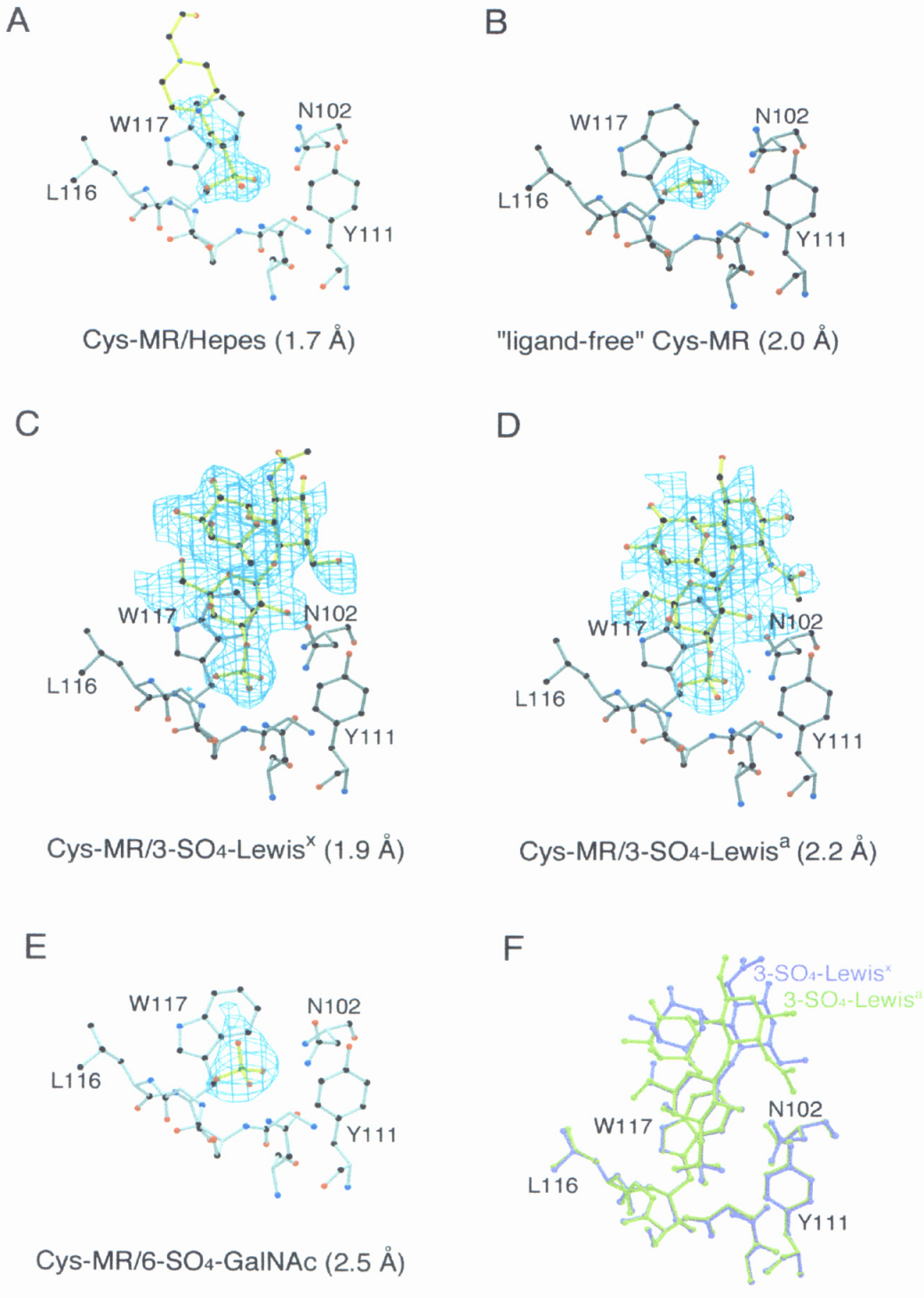


Figure 4. Liu et al.

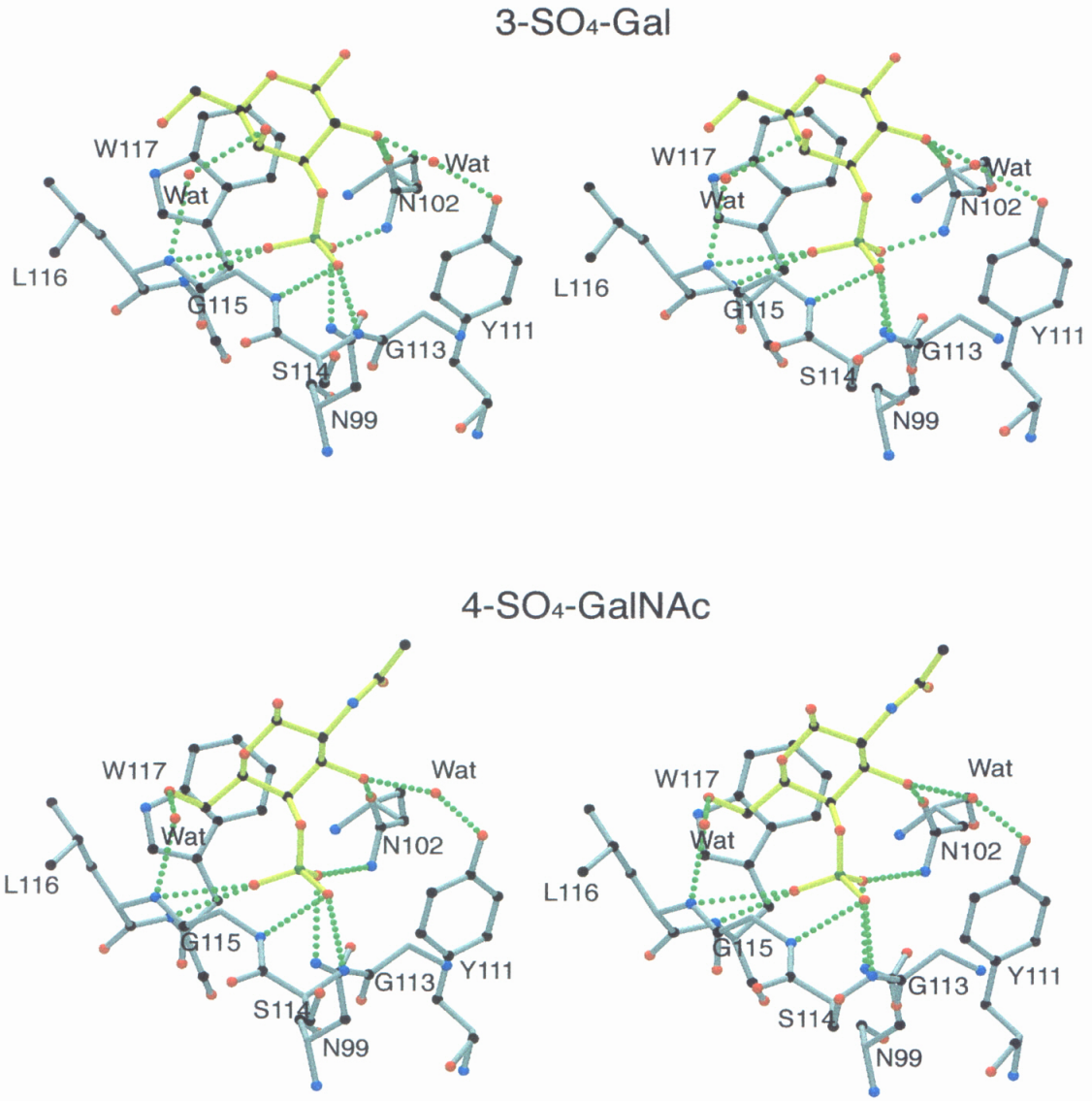


Figure 5. Liu et al.

Appendix I

Expression and Crystallization of Mouse NK Cell Inhibitory Receptors of the Ly49 Family

Abstract

This appendix describes the expression of the mouse NK cell receptor Ly49 in two expression systems: Chinese hamster ovary (CHO) cells and baculovirus infected insect cells. Four soluble forms of Ly49A have been expressed in the baculovirus system and three forms (one for Ly49A and two for Ly49F) were expressed in CHO cells. Among these forms, the baculovirus-derived Ly49A/CRD3 and Ly49A/EC, and CHO cell-derived Ly49F/CRD4 and Ly49A/CRD4, were produced in large scale, purified and screened in crystallization trials. Biochemical characterization of CHO cell-derived Ly49A/CRD4 and Ly49F/CRD4 indicated that they exist as homodimers in solution, an observation consistent with the characterization of Ly49 receptors at the cell surface. In the attempt to obtain large single Ly49 crystals, we have extensively explored crystallization conditions and applied seeding methods. Micro-crystals were obtained for the baculovirus-derived Ly49A/CRD3 and CHO cell-derived Ly49F/CRD4 proteins, and macro-crystals were obtained for the CHO cell-derived Ly49A/CRD4 protein.

Introduction

Natural Killer (NK) cells are a lineage of lymphocytes, distinct from T and B lymphocytes, which play a prominent role in the host's innate immune responses (Janeway and Travers, 1997). They can lyse a variety of target cells without prior immunization (Trinchieri, 1989). Recognition of major histocompatibility complexes (MHC) class I molecules on target cells plays a key role in the regulation of NK cell activity. Self-MHC class I molecules expressed by target cells can inhibit NK cell attack. However, target cells having an altered or absent expression of "self" MHC class I molecules on their cell surface can be destroyed by NK cells (Ljunggren and Karre, 1990).

In the mouse, NK cell receptors specific for MHC class I molecules include the Ly49 and CD94/NKG2 receptor families. The Ly49 gene family, which contains at least nine members (Ly49A-I), is part of the murine NK complex located on mouse chromosome 6 (Brown et al., 1997; McQueen et al., 1998; Yokoyama, 1995). Members of the Ly49 family are expressed on different but overlapping subsets of NK cells, so that a single NK cell can have several Ly49 receptors at the cell surface (Raulet et al., 1997). Each Ly49 receptor can bind different sets of classical MHC class I molecules. For example, Ly49A binds H-2D^{d} and H-2D^{k}, and Ly49C binds H-2K^{b}, H-2K^{d}, H-2D^{d} and H-2^{s} encoded MHC class I molecules (Hanke et al., 1999; Takei et al., 1997).

Ly49 receptors are disulfide-linked homodimers with type II membrane orientation. Each receptor contains two extracellular regions: 1) a C-type lectin-like domain called a carbohydrate recognition domain (CRD) that is homologous to the superfamily of Ca^{2+}-dependent lectins, which have known carbohydrate binding properties (Drickamer and Taylor, 1993); 2) a stalk region that is involved in the dimerization of Ly49 receptors

(Yokoyama and Seaman, 1993). The binding ability of Ly49 to MHC class I molecules is probably conferred by its CRD and part of the stalk region (Brennan et al., 1996).

In this study, I have chosen Ly49A as a prototype molecule to express and crystallize, since Ly49A is the best-characterized member in the Ly49 family. Four soluble forms of Ly49A were expressed in the baculovirus expression system and one form in Chinese hamster ovary (CHO) cells. In addition, two soluble forms of Ly49F were expressed in CHO cells. Biochemical characterization of CHO cell-derived Ly49A/CRD4 and Ly49F/CRD4 indicates that they exist as homodimers in solution, an observation consistent with their existence at the cell surface. Four Ly49 proteins (baculovirus-derived Ly49A/EC and Ly49A/CRD3 and CHO cell-derived Ly49A/CRD4 and Ly49F/CRD4) were expressed in large scale, purified, and put into crystallization trials. Micro-crystals were obtained for the baculovirus-derived Ly49A/CRD3 and the CHO cell-derived Ly49F/CRD4 proteins, and macro-crystals were obtained for the CHO cell-derived Ly49A/CRD4 protein.

Materials and Methods

Overview of Ly49 constructs

The Ly49 receptor is a type II transmembrane molecule, with its N-terminus inside the membrane and its C-terminus outside. Unlike type I transmembrane molecules, Ly49 lacks the secretion signal peptide that is critical for protein secretion. To overcome this problem, I fused a secretion signal peptide to the N-terminus of Ly49 constructs and successfully expressed soluble versions of Ly49 molecules. In this study, I have expressed Ly49A and Ly49F proteins, and focused primarily on the expression of Ly49A. In order to obtain soluble Ly49 proteins that can be crystallized, I have built the following constructs (Figure 1):

1) **EC** – the extracellular (EC) domain. It contains 198 residues, from residue 66 to residue 263.
2) **CRD2** – a truncated C-terminal domain homologous to the carbohydrate recognition domain (CRD) of rat mannose binding protein (MBP-A), which has a known three-dimensional structure (Weis et al., 1992). The CRD2 region shows weak sequence identity (about 20%) to the CRD of MBP-A. It contains a total of 141 residues, from residue 123 to residue 263.
3) **CRD3** – a longer version than CRD2, with a total of 148 residues (from residue 116 to residue 263). It contains the CRD2 and part of the stalk region that is probably involved in ligand binding (Brennan et al., 1996).
4) **CRD4** - the longest CRD version, with a total of 161 residues (from residue 103 to residue 263). It contains the CRD3 and an extra part of the stalk region, including a cysteine residue that is involved in covalent dimerization of Ly49 receptors.

Construction of expression vectors to express truncated forms of Ly49 receptor

1) For expression in the baculovirus expression system

Ly49A gene was kindly provided by David Raulet at UC Berkeley. Four different *Ly49A* constructs were generated using PCR and are designated as CRD2, CRD3, CRD4, and EC (Figure 1). Oligonucleotide primers introducing a 5' EcoR I site, a 3' Not I site, and a stop codon preceding the Not I site for PCR amplifications were:

for CRD2, 5' primer, 5'- AGTCAGGAATTCAAAACCAAGACTGTTTTAGAT,

3' primer, 5'- GAAGCTTGCGGCCGCTCAATGAGGGAATTTATC;

for CRD3, 5' primer, 5' - GGATCCGCATCATCATCATCATCACAGCAGCGGCATCGAGGGACGTG GTGAATTCGATCAGAACAGATTGTAT,

3' primer, the same as for Ly49A/CRD2;

for CRD4, 5' primer, 5'- AGTCAGGAATTCAAGTCTATAGAGTGTGATCTT,

3' primer, the same as for Ly49A/CRD2;

for EC, 5' primer, 5'- GAATTCGAAACCAAGACGAAAATTTTTCAGTAT ,

3' primer, the same as for Ly49A/CRD2.

The resulting PCR products were digested by EcoR I and Not I, and then subcloned into a modified vector pAcGP67A (Invitrogen), which contains codons encoding a strong, insect cell specific, secretion signal peptide, a 6 x His tag, and a Xa cleavage sequence 5' to Ly49 constructs (Figure 2A).

2) For expression in mammalian cells (CHO cells)

Ly49A constructs

Ly49A/CRD4 construct was fused 3' to a gene fragment encoding the hydrophobic leader peptide from neonatal rat Fc receptor (FcRn) (Simister and Rees, 1985) by bridge PCR. Oligonucleotide primers used for the construct were:

5' primer, 5'- CCTCGAGCCACCATGGGGATGTCCCAGCCC,

3' primer, 5'- GAATTCGAAACCAAGACGAAAATTTTTCAGTAT,

bridge primer, 5' -TCAGACCTGGGGAGCGGAGCCCCGTAAGTCTATAGAGTGTGATCTTCTGG.

The bridge primer introduced codons encoding AEPR immediately at the N-terminus of Ly49A/CRD4 protein (Figure 2B).

A 5' Xho I site, a 3' Not I site and a stop codon preceding the Not I site were introduced by PCR. The PCR product was digested by Not I and Xho I, and then subcloned into expression vector PBJ5/GS that was digested by the same enzymes (Chapman and Bjorkman, 1988).

Ly49F constructs

PCR was used to generate two constructs for Ly49F gene (provided by David Raulet, UC Berkeley), designated as Ly49F/CRD2 and Ly49F/CRD4. Each construct encodes an N-terminal 6 x His tag followed by Ly49F/CRD2 or Ly49F/CRD4, which has similar length to Ly49A/CRD2 or Ly49A/CRD4, respectively. Oligonucleotide primers that produced a 5' Xho I site, a 3' Not I site and a stop codon preceding the Not I site for PCR amplifications were:

for Ly49F/CRD2,

5' primer, 5'- ACTCTCGAGCATCATCATCATCACAGCAGCGGCATCGAGGGACGTGGTACAAAGACAGATTTAGATTCCTCA,

3' primer, 5'- ACTGAAGCTTGCGGCCGCTCATTAATGAGGAAATTTATCCAG;

for Ly49F/CRD4,

5' primer, 5'-ATCGCGCTCGAGCATCATCATCATCACAGCAGCGGCATCGAGGGACGTGGTACAGACTCTAGGCCAGGCAATGAACTT,

3' primer, the same as for Ly49F/CRD2.

PCR products were digested by Xho I and Not I, and then subcloned into the unique the Xho I and Not I sites of a modified PBJ5/GS vector (generated by Watson Shen and Lance Martin). The modified vector contains codons encoding the secretion signal peptide of human IgG1 5' to the Xho I site (Figure 2C). It introduces extra codons encoding amino acid sequence LE at the amino terminal to the subcloned constructs (Figure 2C).

Expression and purification of recombinant Ly49 proteins

1) Baculovirus-derived Ly49A proteins

Recombinant viruses were generated by co-transfection of the expression vectors with linearized viral DNA (Baculogold; Pharmingen) (by Dr. Peter Snow, at Caltech Protein Expression Center). Soluble forms of Ly49A protein were secreted into supernatants by baculovirus-infected H5 cells. Recombinant proteins were separated from cell supernatants using Ni-NTA chromatography (Ni-NTA superflow column; Qiagen), and further purified by ion-exchange chromatography using an FPLC Mono-S column (Pharmacia) followed by size-exclusion chromatography using an FPLC Superdex-200 column (Pharmacia).

2) CHO cell-derived Ly49 proteins

Expression

The Ly49A/CRD4 expression vector was transfected into Chinese hamster ovary (CHO) cells by lipofectin transfection (GIBCO, BRL). Transfected cells, which should be resistant to 100 μM methionine sulfoximine (MSX) (Sigma), were selected according to a protocol established by Celltech (Glover, 1987). Transfected CHO cell clones were maintained in glutamine-free αMEM medium (Irvine Scientific) that contains 5% dialyzed fetal bovine serum (DFCS) (GIBCO, BRL). After being metabolically labeled with [^{35}S] methionine and [^{35}S] cysteine (ICN), cell clones secreting Ly49A/CRD4 protein were

identified by immunoprecipitation of supernatants using A1 monoclonal antibody (mAb) against Ly49A receptor (Nagasawa et al., 1987). Positive clones were picked if immunoprecipitation yielded a labeled protein with molecular weight of 22kD for Ly49A/CRD4.

The Ly49F/CRD2 and Ly49F/CRD4 expression vectors were transfected into CHO cells using the same procedure mentioned above. For selecting cell clones secreting Ly49F/CRD2 or Ly49F/CRD4 proteins, anti-His mAb against 6 x His tag (Qiagen) was used for immunoprecipitation. Positive clones were picked if immunoprecipitation yielded a labeled protein with molecular weight of 20kD for Ly49F/CRD2 or 22kD for Ly49F/CRD4.

Purification

Cell lines secreting Ly49A/CRD4, Ly49F/CRD2, and Ly49F/CRD4 proteins were expanded to confluence in twenty 15cm plates. Supernatants were collected every two days for 4-8 weeks. Different purification schemes were used for Ly49A and Ly49F proteins.

Cell supernatants containing Ly49A/CRD4 protein were filtered to remove cell debris, and then were added 1/30 volume of 1M MES buffer pH 5.0 to make final pH between 5.0 to 5.5. The buffered supernatants were passed over an ion exchange column, SP column (Pharmacia). The column was washed using a solution containing 20mM MES pH 5.5 and 20mM NaCl, and eluted using a solution containing 20mM Tris pH 7.4 and 1M NaCl. The eluted Ly49A/CRD4 protein was further purified by ion exchange chromatography using an FPLC Mono-S column followed by size-exclusion chromatography using an FPLC Superdex-200 column. Approximately 5 mg of purified Ly49A/CRD4 protein was obtained per liter of cell supernatants.

Cell supernatants containing Ly49F/CRD2 or Ly49F/CRD4 were thoroughly dialyzed against TBS (20mM Tris pH 7.4 and 150mM NaCl). Ly49F/CRD2 or Ly49F/CRD4 protein was purified using Ni-NTA followed by size-exclusion chromatography using an FPLC superdex-200 column. Approximately 1mg of Ly49F/CRD4 or 0.2 mg of Ly49F/CRD2 protein was obtained per liter of supernatants.

Deglycosylation of Ly49A proteins by treatment with glycosidases

Purified soluble Ly49A proteins were treated with glycosidases at 25°C or 4°C in their reaction conditions, according to the manufacturer's protocols. The following glycosidases were tested to deglycosylate Ly49A proteins: endoglycosidase H, endoglycosidase F, N-glycosidase F, neuraminidase (Boehringer Mannheim), and mannosidase (Sigma). Glycosidase-treated proteins were analyzed by SDS-PAGE to monitor the deglycosylation. In this study, I have deglycosylated three proteins for crystallization purpose, including the CHO cell-derived Ly49A/CRD4, baculovirus-derived Ly49A/CRD3 and Ly49A/EC proteins.

Crystallization of soluble Ly49 proteins

The purified Ly49A or Ly49F proteins were changed into 10mM HEPES (pH 7.4) and concentrated (15 mg/ml for the baculovirus-derived Ly49A/CRD3, 15 mg/ml for the baculovirus-derived Ly49A/EC, 40 mg/ml for the CHO-derived Ly49A/CRD4, and 12 mg/ml for the CHO derived Ly49F/CRD4). Proteins were mixed with precipitating solutions from Hampton screen kits (crystal screen I and II, Hampton Research) at 1:1 ratio in hanging-drops. Crystals were grown by vapor diffusion method at 22°C and 4°C. If micro-

crystals were grown from certain solutions, those solutions were screened and optimized to grow larger crystals (Abelson and Simon, 1997). Microseeding and macroseeding methods were also applied to grow large crystals.

Results and discussion

1) CHO cell-derived Ly49 proteins

A gene fragment encoding the secretion signal sequence of rat FcRn or the secretion signal sequence of human IgG1 was fused 5' to Ly49A or Ly49F constructs, respectively. The resulting constructs were subcloned into the expression vector PBJ5/GS. CHO cells were transfected, and clones secreting recombinant proteins were selected (see *Materials and Methods*). Stable transformant CHO cell lines were identified by immunoprecipitating metabolically labeled supernatants with A1 mAb for Ly49A/CRD4 or anti-His mAb for the His-tagged Ly49F proteins. The highest expressing CHO clones were selected and expanded. Cell supernatants were collected for purification.

a) Purification of Ly49A/CRD4 protein

Supernatants produced by Ly49A/CRD4-secreting cells were passed over an SP column to isolate Ly49A/CRD4 protein from the medium. Ly49A/CRD4 protein was further purified using a Mono-S column followed by a superdex-200 column on FPLC. As shown in its gel filtration profile (Figure 3A), Ly49A/CRD4 protein migrates in a single peak at a position corresponding to 40 kD, suggesting Ly49A/CRD4 exists as homodimers in solution. SDS-PAGE analysis confirmed this suggestion. Under reducing conditions, Ly49A/CRD4 migrates with an apparent molecular mass of 22 kD, but under non-reducing condition, it migrates with molecular mass of 40 kD (Figure 3A). The expressed Ly49A/CRD4 protein is glycosylated (it contains one N-linked glycosylation site). The apparent molecular mass of Ly49A/CRD4 is 22 kD, 3 kD higher than the calculated molecular mass (19 kD). The purified Ly49A/CRD4 was > 95% pure. About 5mg of the purified Ly49A/CRD4 protein was obtained per liter of supernatants.

b) Purification of Ly49F/CRD4

Supernatants from CHO cells secreting Ly49F/CRD2 or Ly49F/CRD4 were dialyzed and then passed over a Ni-NTA column. After this step, yields of Ly49F/CRD2 and Ly49F/CRD4 were assessed. Approximate 0.2 mg of Ly49F/CRD2 or 1 mg of Ly49F/CRD4 was obtained per liter of supernatant. Due to the low yield of Ly49F/CRD2 protein, large scale expression and further purification was not pursued. Ly49F/CRD4 protein was further purified using a superdex-200 column on FPLC. By gel filtration, Ly49F/CRD4 protein migrates to almost the same position as Ly49A/CRD4 (Figure 3B). This result indicates that Ly49F/CRD4 also exists as homodimers in solution. Ly49F receptor is believed to exist as a non-covalent homodimer at the cell surface, while the other Ly49 receptors form disulfide-linked homodimers. The reason is that Ly49F has a mutation (Cys to Ser) on the conserved cysteine residue that forms disulfide bond. This speculation is confirmed by the results of SDS-PAGE analysis of purified Ly49F/CRD4 protein, which reveals one band with molecular mass of 22 kD under non-reducing condition (Figure 3B). Ly49F/CRD4 protein is not glycosylated (it contains no N-linked glycosylation site). Its apparent molecular mass shown by SDS-PAGE is almost the same as the calculated molecular mass (21 kD). Approximately 1mg of the purified Ly49F/CRD4 protein was obtained per liter of supernatants.

2) Expression of soluble forms of Ly49A protein in the baculovirus system

To produce large amounts soluble Ly49A protein for crystallization, I have built four cDNA constructs encoding different extracellular regions of Ly49A – designated as Ly49A/EC, Ly49A/CRD2, Ly49A/CRD3, and Ly49A/CRD4 (see *Materials and Methods*).

To express soluble Ly49A proteins, a gene fragment encoding a strong secretion signal peptide, a 6 x His tag and a Xa cleavage sequence was fused 5' to the Ly49A constructs. The resulting expression constructs were subcloned into a transfer vector pAcGP67A, and the constructed vectors were used to make recombinant viruses. Insect H5 cells were infected by the recombinant viruses to produce soluble Ly49A proteins. Supernatants of infected cells were passed over a Ni-NTA column to separate Ly49A proteins from the medium. The eluted Ly49A proteins were analyzed by SDS-PAGE. Results for the four soluble forms of Ly49A were summarized in Table 1. Ly49A/CRD2 and Ly49A/CRD4 have low estimated yields and were not pursued for large scale expression and further purification. The other forms, Ly49A/CRD3 and Ly49A/EC, have good yields (2-3 mg for Ly49A/CRD3 or 3-5 mg for Ly49A/EC per liter of supernatants). They were expressed in large scale and further purified.

1) Purification of Ly49A/CRD3

Ly49A/CRD3 protein was isolated from supernatants using a Ni-NTA column, and was further purified using a Mono-S column followed by a Superdex-75 column on FPLC. As shown by its gel filtration profile (Figure 4A), Ly49A/CRD3 protein migrates to a position corresponding to 30 kD. Using SDS-PAGE analysis, Ly49A/CRD3 migrates as a single band with apparent molecular mass of 21 kD under reducing and non-reducing conditions. These data indicate that Ly49A/CRD3 exists as monomers in solution. Ly49A/CRD3 has a shorter stalk region (7 amino acids shorter) than Ly49A/CRD4 (characterized as homodimers in solution), so that it lacks the cysteine residue involved in covalent dimerization. Finding that Ly49A/CRD3 is a monomer indicates the important role of the stalk region and the cysteine residue in dimerization. Ly49A/CRD3 is a glycoprotein

(it has one N-linked glycosylation site.). The apparent molecular mass of Ly49A/CRD3 is 21 kD, 2 kD higher than the calculated molecular mass (19 kD). Approximately 3 mg of the purified Ly49A/CRD3 protein was obtained per liter of supernatants.

2) Purification of Ly49A/EC

The purification of Ly49A/EC followed the same procedure as that of Ly49A/CRD3. As shown by its gel filtration profile (Figure 4B), Ly49A/EC migrates as a single peak at a position corresponding to 60 kD, indicating that Ly49A/EC exists as a homodimer in solution. Under non-reducing condition, SDS-PAGE analysis of purified Ly49A/EC revealed 2 bands, one having an apparent molecular mass of 60 kD, corresponding to Ly49A/EC dimer and the other one having an apparent molecular mass of 31 kD, corresponding to Ly49A/EC monomer. Under reducing condition, SDS-PAGE analysis revealed only one band, corresponding to the molecular mass of Ly49A/EC monomer. These data indicate that the baculovirus-derived Ly49A/EC protein is a mixture of covalently linked and non-covalently linked homodimers. The expressed Ly49A/EC is glycosylated, containing three N-linked glycosylation sites. Its apparent molecular mass revealed by SDS-PAGE analysis is 31 kD, showing an excess of 8 kD for carbohydrates compared with the predicted molecular mass of 23 kD. About 5mg of the purified Ly49A/EC was obtained per liter of supernatants.

3) Deglycosylation

In attempts to remove carbohydrates on the recombinant Ly49 proteins, I have tested different glycosidases, including endoglycosidase H, endoglycosidase F, N-glycosidase F, neuraminidase, and mannosidase. For the CHO cell-derived Ly49A/CRD4 protein, N-glycosidase F can efficiently remove the carbohydrates and reduce the apparent molecular

mass to 18 kD, as shown by SDS-PAGE analysis (Figure 5A). The baculovirus-derived Ly49A/CRD3 and Ly49A/EC proteins bear carbohydrates with high mannose structures that are modified by insect cells. Mannosidase was used for deglycosylating these proteins. It reduced the apparent molecular mass about 1 kD for Ly49A/CRD3 and 1 kD for Ly49A/EC (Figure 5B). The deglycosylated proteins were repurified using size exclusion chromatography and placed into crystallization trials.

4) Crystallization

The hanging drop method was used to screen suitable crystallization conditions for soluble Ly49 proteins. In this study, a drop (1-2 microliters) of Ly49 protein was mixed with the trial precipitant solutions in equal volumes, and equilibrated with 1ml of the original precipitant solution. Initial attempts to crystallize Ly49 proteins were implemented using the precipitating solutions from Hampton's crystal screen I and II kits at 22°C and at 4°C. Conditions that produced microcrystals are summarized in Table 2. Since the average intensity of the X-ray diffraction pattern is proportional to the size of crystal, it is very important to grow crystals with a substantial size.

To obtain conditions for growing large crystals, I have screened different parameters, including pH, the protein concentrations, the precipitant concentrations, different precipitants, additives, and temperatures. Also, microseeding and macroseeding were used to improve the size and quality of crystals. Unfortunately, my efforts to obtain large single crystals were not successful. The biggest crystals are the crystals of CHO cell-derived Ly49A/CRD4 protein grown in a solution containing 0.1M Cacodylate pH 6.5, 3.8M NH_4Ac, and 10% isopropanol at 22°C (Figure 6).

In a second approach, I have used glycosidases to deglycosylate Ly49A proteins (N-glycosidase F for CHO cell-derived Ly49A/CRD4 and mannosidase for baculovirus-derived Ly49A/CRD3 and Ly49A/EC). The deglycosylated proteins were repurified using size exclusion chromatography and put into crystallization trials. However, treatment with glycosidases did not improve the size and quality of crystals significantly.

References:

Abelson, J. N., and Simon, M. I. (1997). Methods in Enzymology. In Macromolecular crystallography, C. W. J. Carter and R. M. Sweet, eds.(Academic press), pp. 23-59.

Brennan, J., Mahon, G., Mager, D. L., Jefferies, W. A., and Takei, F. (1996). Recognition of class I major histocompatibility complex molecules by Ly-49: specificities and domain interactions. J. Exp. Med. *183*, 1553-1559.

Brown, M. G., Fulmek, S., Matsumoto, K., Cho, R., Lyons, P. A., Levy, E. R., Scalzo, A. A., and Yokoyama, W. M. (1997). A 2-MB YAC contig and physical map of the natural killer gene complex on mouse chromosome 6. Genomics *42*, 16–25.

Chapman, T. L., and Bjorkman, P. J. (1988). Characterization of a murine cytomegalovirus class I major histocompatibility complex (MHC) homolog: comparison to MHC molecules and to the human cytomegalovirus MHC homolog. J. Virol. *72*, 460-466.

Drickamer, K., and Taylor, M. E. (1993). Biology of animal lectins. Annu. Rev. Cell. Biol. *9*, 237-264.

Glover, D. M. (1987). DNA cloning: a practical approach. In The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells, C. R. Bebbington and C. C. G. Hentschel, eds. (Oxford, United Kingdom: IRL press).

Hanke, T., Takizawa, H., McMahon, C. W., Busch, D. H., Pamer, E. G., Miller, J. D., Altman, J. D., Liu, Y., Cado, D., Lemonnier, F. A., Bjorkman, P. J., and Raulet, D. H. (1999). Direct assessment of MHC class I binding by seven Ly49 inhibitory NK cell receptors. Immunity *11*, 67-77.

Janeway, C., and Travers, P. (1997). Immunobiology: The Immune System in Health and Disease, the 3rd Edition (Garland Pub).

Ljunggren, H. G., and Karre, K. (1990). In search of the 'missing self': MHC molecules and NK cell recognition. Immunol. Today *11*, 237-244.

McQueen, K. L., Freeman, J. D., Takei, F., and Mager, D. (1998). Localization of five new Ly49 genes, including three closely related to Ly49C. Immunogenetics *48*, 174–183.

Nagasawa, R., Gross, J., Kanagawa, O., Townsend, K., Lanier, L. L., Chiller, J., and Allison, J. P. (1987). Identification of a novel T cell surface disulfide-bonded dimer distinct from the / antigen receptor. J. Immunol. *138*, 815-824.

Raulet, D. H., Held, W., Correa, I., Dorfman, J., Wu, M.-F., and Corral, L. (1997). Specificity, tolerance and developmental regulation of natural killer cells defined by expression of class I-specific Ly49 receptors. Immunol. Rev. *155*, 41–52.

Simister, N. E., and Rees, A. R. (1985). Isolation and characterization of an Fc receptor from neonatal rat small intestine. Eur. J. Immunol. *15*, 733-738.

Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M., and Arai, N. (1988). SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. Mol. Cell. Biol. *8*, 466-472.

Takei, F., Brennan, J., and Mager, D. L. (1997). The Ly49 family: genes proteins and recognition of class I MHC. Immunol. Rev. *155*, 67–77.

Trinchieri, G. (1989). Biology of natural killer cells. Adv. Immunol. *47*, 187–376.

Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992). Structure of a C-type mannose-binding protein complexed with an oligosaccharide. Nature *360*, 127-134.

Yokoyama, W. M. (1995). Natural killer cell receptors specific for major histocompatibility complex class I molecules. Proc. Natl. Acad. Sci. USA *92*, 3081-3085.

Yokoyama, W. M., and Seaman, W. E. (1993). The Ly-49 and NKR-P1 gene families encoding lectin-like receptors on natural killer cells: the NK gene complex. Annu. Rev. Immunol. *11*, 613-635.

Figure 1. The Ly49A constructs

Four Ly49A constructs designated as EC, CRD2, CRD3, and CRD4 are shown above (described in *Materials and Methods*). Potential N-linked glycosylation sites are indicated as "Y". The cysteine residue in the stalk region that forms disulfide bond is also indicated as "-S".

Figure 2. The Ly49 expression constructs for expression in the baculovirus system and in CHO cells

(A) For expression in the baculovirus system, the expression construct of Ly49A encodes (from the N-terminus) the secretion signal sequence of gp67 protein, a 6 x His tag, the Xa cleavage sequence, and the Ly49A construct. Expression is under the control of polyhedrin promoter.

(B) For CHO cell expression, the expression construct of Ly49A/CRD4 encodes (from the N-terminus) the secretion signal sequence of rat FcRn, a linker with the sequence of AEPR, and the Ly49A/CRD4 construct. The CHO cell expression is under the control of SRα promoter (Takebe et al., 1988).

(C) For CHO cell expression, the expression construct of Ly49F encodes (from the N-terminus) the secretion signal sequence of human IgG1, a linker with the sequence of LE, a 6x His tag, and the Ly49F/CRD2 or Ly49F/CRD4 construct. The CHO cell expression is under the control of SRα promoter.

Figure 3. Gel filtration and SDS-PAGE analysis of the CHO cell-derived Ly49A/CRD4 and Ly49F/CRD4 proteins

(A) The purified Ly49A/CRD4 protein was run on the FPLC Superdex-200 column. The recovered peak was analyzed by SDS-PAGE under reducing (R) and non-reducing (NR) conditions as shown in the inset.

(B) The purified Ly49F/CRD4 protein was run on the FPLC Superdex-200 column. The recovered peak was analyzed by SDS-PAGE under non-reducing (NR) condition as shown in the inset.

Figure 4. Gel filtration and SDS-PAGE analysis of baculovirus-derived Ly49A/CRD3 and Ly49A/EC proteins

(A) The purified Ly49A/CRD3 protein was run on the FPLC Superdex-75 column. The recovered peak was analyzed by SDS-PAGE under reducing (R) and non-reducing (NR) conditions as shown in the inset.

(B) The purified Ly49A/EC protein was run on the FPLC Superdex-75 column. The recovered peak was analyzed by SDS-PAGE under reducing (R) and non-reducing (NR) conditions as shown in the inset.

Figure 5. Deglycosylation of Ly49A proteins

(A) The purified baculovirus-derived Ly49A/EC and Ly49A/CRD3 proteins (lane 1 and 3) were deglycosylated with mannosidase (lane 2 and 4).

(B) The purified CHO cell-derived Ly49A/CRD4 (lane 1) was deglycosylated with N-glycosidase F (lane 2).

Figure 6. Crystals of the CHO-cell derived Ly49A/CRD4 protein

Crystals were grown by the vapor diffusion method and photographed in the droplet.

Table 1. Soluble forms of Ly49A expressed using the baculovirus system

protein	apparent M.W. on SDS-PAGE	calculated M.W.	protein form (SDS-PAGE)	yield (mgs/L)
EC	31 kD	23.0 kD	monomers/dimers	3-5
CRD2	18 kD	18.2 kD	monomers	0.1-0.2
CRD3	21 kD	19.1 kD	monomers	2-3
CRD4	23 kD	21.2 kD	monomers/dimers	0.5-1

Four soluble forms of Ly49A were expressed in the baculovirus system, and purified using Ni-NTA chromatography. The purified proteins were analyzed by SDS-PAGE. Yield of each form was estimated.

Table 2. Crystallization of soluble Ly49 proteins

protein	original conditions	improved conditions	crystal morphology
Ly49A/CRD3	0.1M HEPES pH7.5 20% PEG 4000 10% isopropanol	0.1M Tris pH 8.5 20% PEG 4000 14% isopropanol	needles
Ly49A/CRD4	2.0 M $(NH_4)_2SO_4$ 5% isopropanol	0.1M Cacodylate pH 6.5 3.8M NH_4Ac 10% isopropanol	Square plates (0.1mm x 0.1mm)
Ly49A/CRD4	0.1M MES pH6.5 1.6 M $(NH_4)_2SO_4$ 10% dioxane		
Ly49A/CRD4	0.1M HEPES pH7.5 0.2M $MgCl_2$ 30% isopropanol (4°C)	0.1M HEPES pH7.5 0.3M $MgCl_2$ 25% isopropanol (4°C)	rods and plates
Ly49F/CRD4	0.1M NaAc pH4.0 0.2 M $(NH_4)_2SO_4$ 20% PEG 4000	0.1M NaFormate pH3.8 0.2 M $(NH_4)_2SO_4$ 16% PEG 4000	needles

CHO cell-derived Ly49A/CRD4 and Ly49F/CRD4 and baculovirus-derived Ly49A/CRD3 and Ly49A/EC proteins were placed into crystallization trials. The original and improved conditions that produced crystals were summarized.

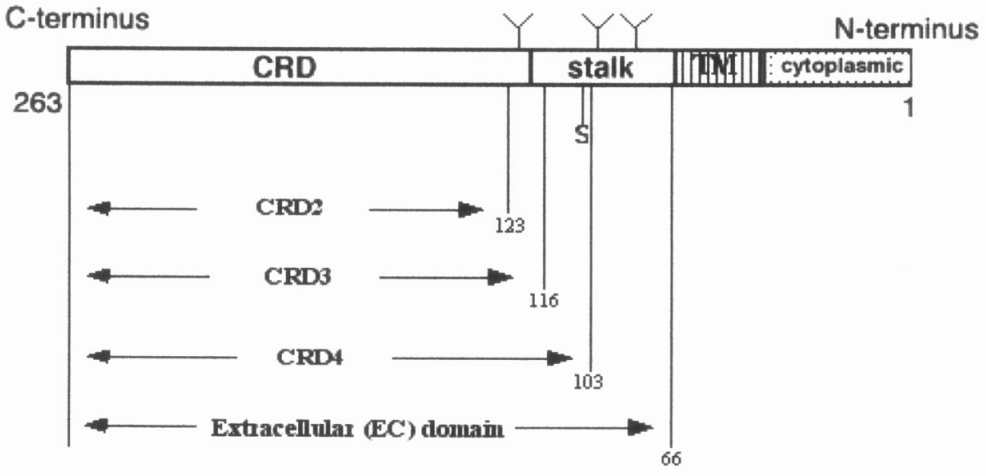


Figure 1. The Ly49A constructs

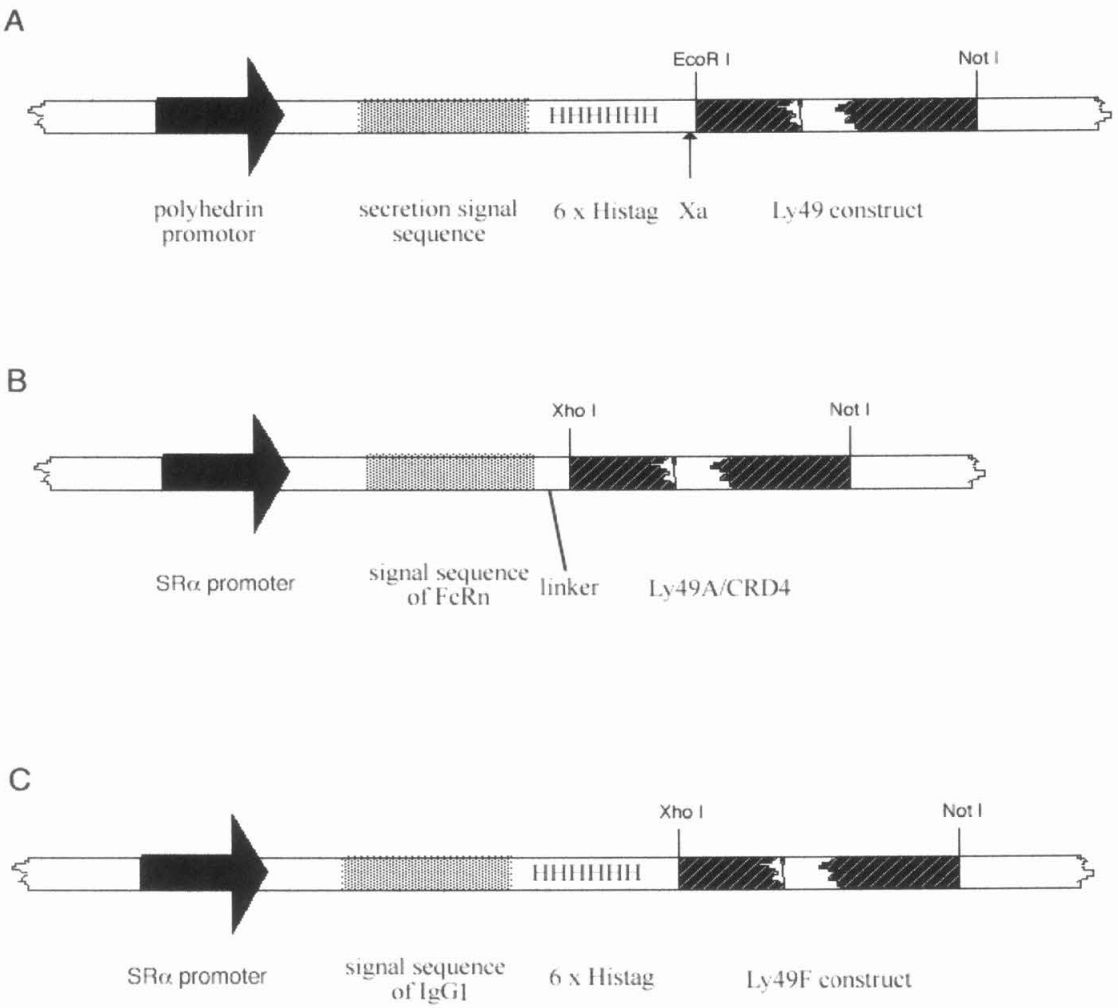


Figure 2. The Ly49 expression constructs for expression in baculovirus system and in CHO cells

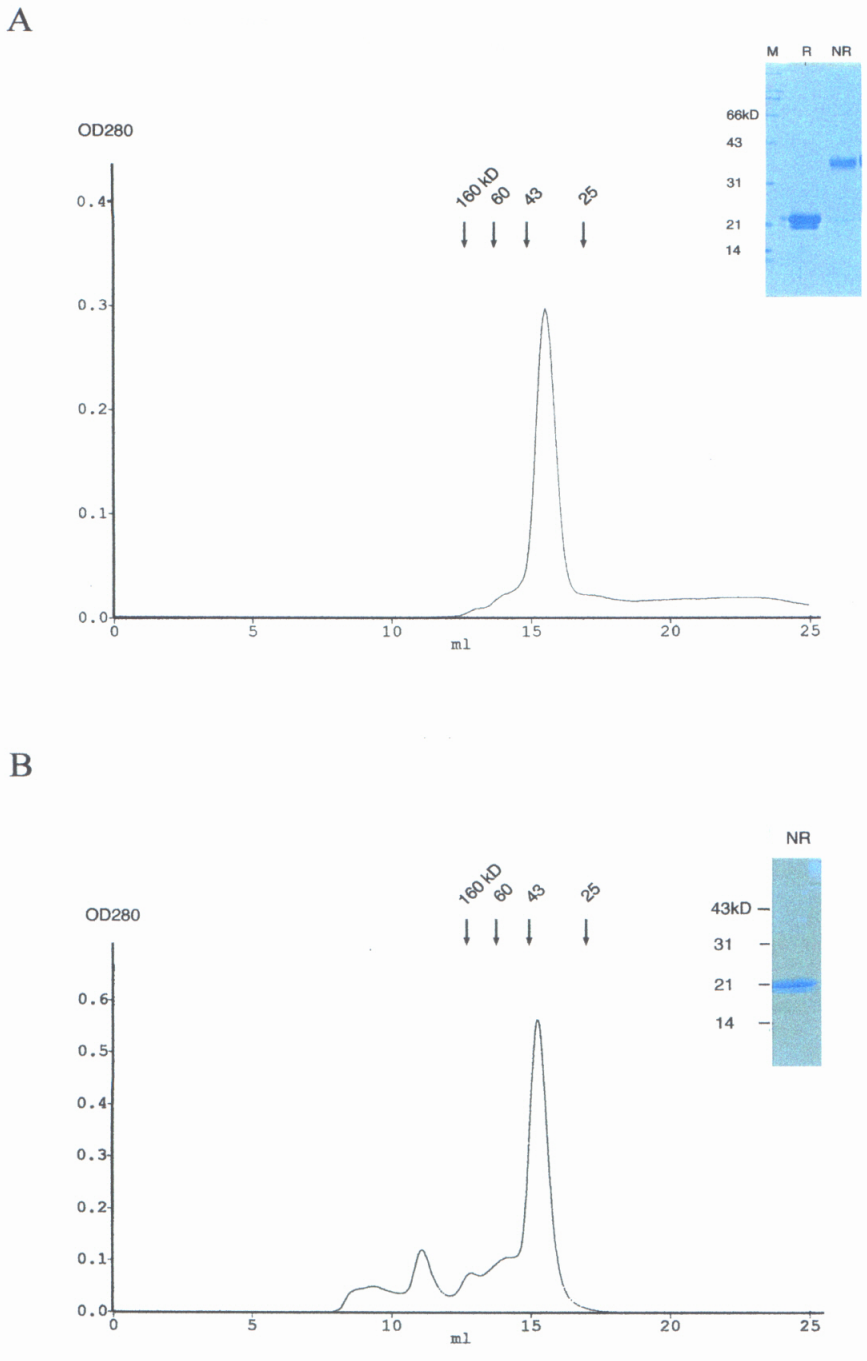


Figure 3. Gel filtration and SDS-PAGE analysis of the CHO cell-derived Ly49A/CRD4 (A) and Ly49F/CRD4 (B)

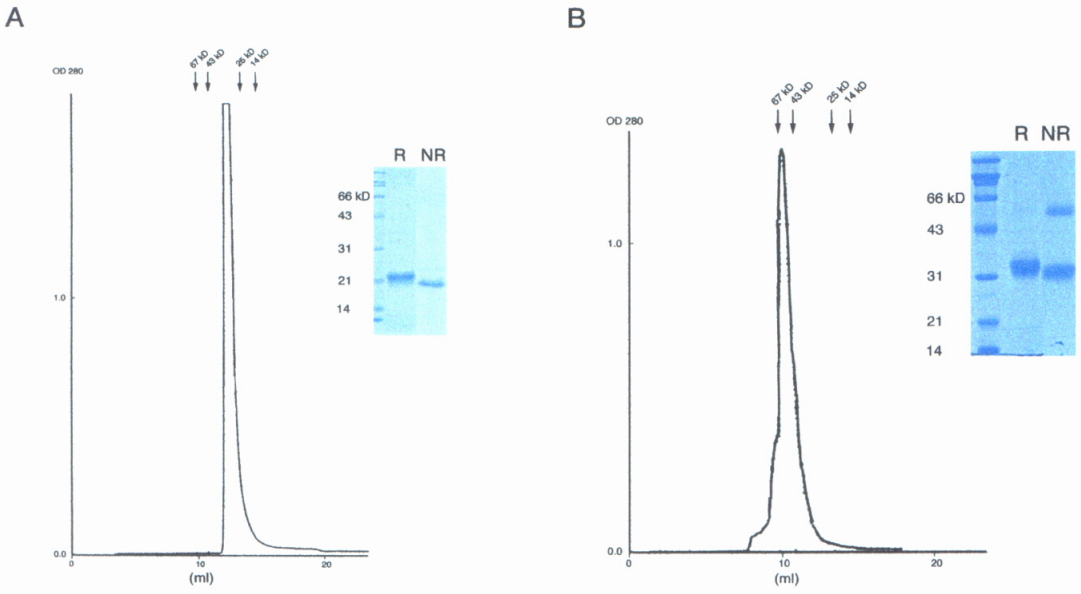


Figure 4. Gel filtration and SDS-PAGE analysis of the baculovirus derived Ly49A/CRD3 (A) and Ly49A/EC (B)

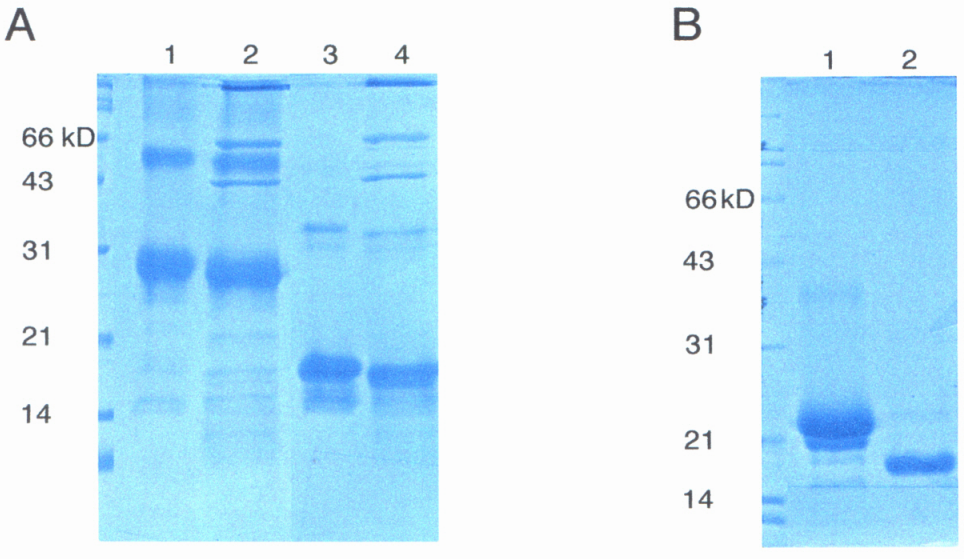


Figure 5. Deglycosylation of Ly49A proteins

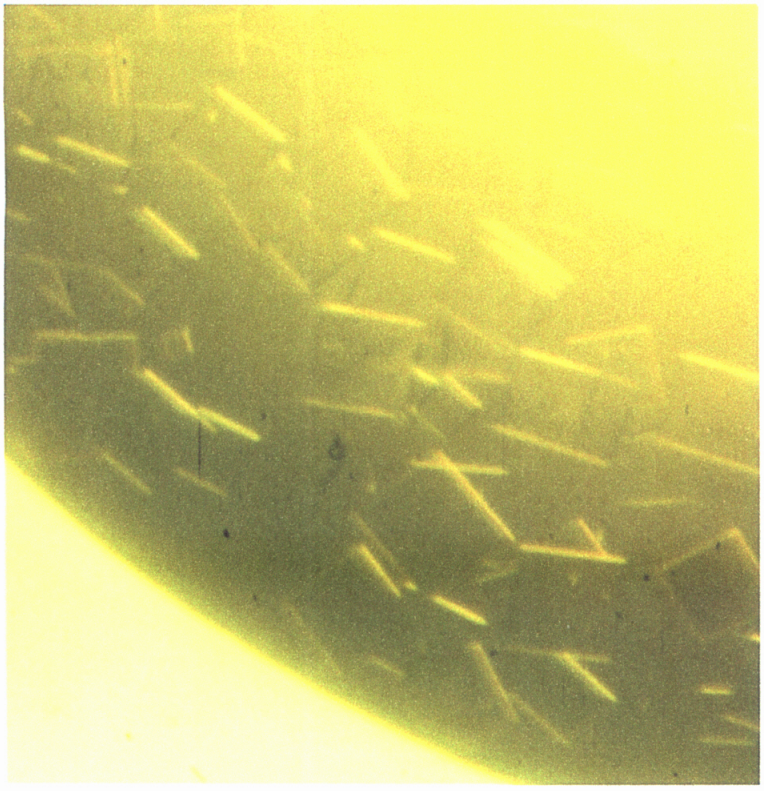


Figure 6. CHO cell-derived Ly49A/CRD4 crystals

Appendix II

Direct Assessment of MHC Class I Binding by Seven Ly49 Inhibitory NK Cell Receptors

Immunity, Vol. 11, 67–77, July, 1999, Copyright ©1999 by Cell Press

Direct Assessment of MHC Class I Binding by Seven Ly49 Inhibitory NK Cell Receptors

Thomas Hanke,[1,7] Hisao Takizawa,[1,8] Christopher W. McMahon,[1] Dirk H. Busch,[2] Eric G. Pamer,[2] Joseph D. Miller,[3] John D. Altman,[3] Yang Liu,[4] Dragana Cado,[1] Francois A. Lemonnier,[5] Pamela J. Bjorkman,[4] and David H. Raulet[1,6]

[1]Department of Molecular and Cell Biology
and Cancer Research Laboratory
University of California
Berkeley, California 94720

[2]Sections of Infectious Diseases and Immunobiology
Yale University School of Medicine New Haven,
Connecticut 06520

[3]Emory Vaccine Center
and Department of Microbiology and Immunology
Emory University School of Medicine
Atlanta, Georgia 30322

[4]Division of Biology
and Howard Hughes Medical Institute
California Institute of Technology
Pasadena, California 91125

[5]Unité d'Immunité Cellulaire Antivirale Département
SIDA-Rétrovirus
Institut Pasteur
F-75724 Paris Cedex 15
France

Summary

Mouse NK cells express at least seven inhibitory Ly49 receptors. Here we employ a semiquantitative cell–cell adhesion assay as well as class I/peptide tetramers to provide a comprehensive analysis of specificities of Ly49 receptors for class I MHC molecules in eight MHC haplotypes. Different Ly49 receptors exhibited diverse binding properties. The degree of class I binding was related to the extent of functional inhibition. The tetramer studies demonstrated that neither glycosylation nor coreceptors were necessary for class I binding to Ly49 receptors and uncovered peptide-specific recognition by a Ly49 receptor. The results provide a foundation for interpreting and integrating many existing functional studies as well as for designing tests of NK cell development and self-tolerance.

Introduction

Recognition of MHC class I molecules on target cells is a key event in the regulation of NK cell activity. Class I–specific inhibitory receptors expressed on the surface of NK cells prevent them from lysing target cells expressing a cognate class I ligand. Consequently, cells with extinguished class I expression often become target cells for NK cells (Ljunggren and Karre, 1990).

In the mouse, the only known NK cell receptors for classical class I molecules belong to the Ly49 family (Yokoyama, 1995). Seven well described Ly49 family members containing inhibitory motifs in their cytoplasmic domain (Ly49A, -B, -C, -E, -F,- G2, and -I) are known to be expressed at least at the mRNA level by NK cells from C57BL/6 (B6) mice. Based on functional studies, Ly49A was initially attributed a specificity for D^d but not K^d, L^d, K^b, or D^b (Karlhofer et al., 1992). Direct binding of Ly49A to D^d was demonstrated in a cell–cell adhesion assay (Daniels et al., 1994a; Takei et al., 1997) or by binding of Ly49A-expressing cells to soluble, plate-bound D^d molecules (Kane, 1994). Ly49A also binds to H-2^k cells (Brennan et al., 1996b). Although genetic evidence has implicated D^k as the ligand (Karlhofer et al., 1994), anti-D^k antibodies reportedly failed to block adhesion of Ly49A-transfected cells to H-2^k cells (Brennan et al., 1996b).

Because Ly49G2$^+$ NK cells lysed H-2^d cells more efficiently in the presence of mAbs against D^d or L^d, this receptor was assigned a specificity for D^d and L^d (Mason et al., 1995). In vivo, Ly49G2-expressing NK cells were capable of rejecting bone marrow allografts from H-2^b, but not H-2^d donors, consistent with the notion that Ly49G2 is an inhibitory receptor for H-2^d (Raziuddin et al., 1996). A recent report implicates D^d but not L^d as a ligand for Ly49G2-expressing NK cells in vivo (Johansson et al., 1998). However, direct evidence for a physical interaction between Ly49G2 and D^d has not been reported.

The MHC specificities of the closely related Ly49C and Ly49I receptors (Brennan et al., 1996a) were initially studied in bone marrow graft rejection experiments. It was concluded that Ly49C and/or Ly49I recognize the K^b molecule (Yu et al., 1996). Cell adhesion experiments with Ly49C- and Ly49I-transfected COS-7 cells and MHC disparate tumor cell lines have demonstrated that Ly49C binds to H-2^b, H-2^d, H-2^k, and H-2^s gene products, while Ly49I was not found to bind any of these (Brennan et al., 1996a, 1996b). To date, little is known about the expression and MHC reactivity of any of the other Ly49 receptors with predicted inhibitory function (Ly49B, -E, and -F). Ly49D exhibits activating rather than inhibitory function in response to target cells expressing the D^d class I molecule (George et al., 1999; Nakamura et al., 1999), but no binding data has been reported.

Evidence indicates that Ly49A interacts with the $\alpha 2$ domain of D^d (Daniels et al., 1994a; Sundbäck et al., 1998), with possible contributions from the $\alpha 1$ domain (Matsumoto et al., 1998). The functional binding of Ly49A to D^d showed no peptide specificity, since a variety of peptides that stabilize D^d expression on target cells protect these from lysis by Ly49A$^+$ NK cells (Correa and Raulet, 1995; Orihuela et al., 1996). The role of D^d glycosylation has been investigated based upon the lectin-like structure of Ly49 receptors and the capacity of both Ly49A and Ly49C to bind the sulfated glycan fucoidan

[6]To whom correspondence should be addressed (e-mail: raulet@uclink4.berkeley.edu).

[7]Present address: Institute for Virology and Immunobiology, University of Würzburg, D-97078 Würzburg, Germany.

[8]Present address: Department of Advanced Pharmacology, Otsuka Pharmaceutical Co., Ltd., Tokushima 771-0192, Japan.

(Daniels et al., 1994b; Brennan et al., 1995). Class I molecules on target cells treated with glycosylation inhibitors or glycosidases reportedly failed to inhibit NK cells (Daniels et al., 1994b; Brennan et al., 1995). However, mutation of the N-glycosylation acceptor sites in D^d did not reduce inhibitory NK cell recognition in one study (Matsumoto et al., 1998) and reduced it only partially in another (Lian et al., 1998). Thus, the role of class I–linked glycans in Ly49 interactions is not fully resolved.

Ly49 receptors are expressed on partially overlapping subpopulations of NK cells (reviewed in Raulet et al., 1997). The missing self hypothesis proposes that NK cells should be inhibited by the few self–class I molecules inherited by each individual. It is believed that "education" processes ensure that many or all functional NK cells express self–class I specific receptors (Hoglund et al., 1997; Raulet et al., 1997). A direct assessment of whether each NK cell expresses at least one self–class I–specific receptor will require knowledge of which Ly49 receptors interact with which MHC ligands. Such information may also aid in elucidating the features of class I molecules that are recognized by Ly49 receptors. To date, the specificities of only a few Ly49 receptors have been examined and only for a few MHC haplotypes. Here, we employ two binding assays and a functional assay to examine the specificities of the panel of inhibitory Ly49 receptors for eight MHC haplotypes.

Results

MHC Allele Specificities of Inhibitory Ly49 Molecules

We initially employed a semiquantitative cell adhesion assay to investigate Ly49 specificity. Ly49-transfected COS-7 cells were adhered to ^{3}H-labeled concanavalin A (Con A) blasts from MHC congenic, MHC recombinant, and MHC mutant mice. The use of MHC congenic Con A blasts instead of tumor cells as in earlier studies circumvented the problem of non-MHC variability in adhesive properties among unrelated tumor cell lines.

All Ly49 members were expressed at high levels on the surface of transfected COS-7 cells. The equivalently strong staining intensity of Ly49C, Ly49F, and Ly49I transfectants with the Ly49C, -F, -I, and -H reactive mAb 14B11 (Corral et al., 1999) indicates that these molecules were expressed at comparable levels (Figure 1A). Ly49C and Ly49I transfectants also stained equivalently with the Ly49C- and Ly49I-reactive mAb 5E6 (data not shown). Similarly, Ly49A transfectants stained strongly with mAb JR9–318, Ly49G2 transfectants with mAb 4D11, and Ly49D transfectants with mAb 4E5 (Figure 1A). Since no Ly49B- or Ly49E-reactive mAbs are available, hemagglutinin (HA) epitope tags were added to the C termini of these molecules, and surface expression was monitored with anti-HA mAb. Control experiments with Ly49A showed that the HA tags did not affect the extent of binding in the cell–cell adhesion assay. Ly49B and Ly49E were expressed to a similar degree as the control-tagged Ly49A (Figure 1B).

We readily observed specific binding of most Ly49 transfectants to Con A blasts from one or more of the B10 MHC congenic mice tested (Figure 2). In cases where Ly49 transfectants bound to Con A blasts from $H\text{-}2^{b,\ d,\ or\ k}$ mice, no binding was observed with class

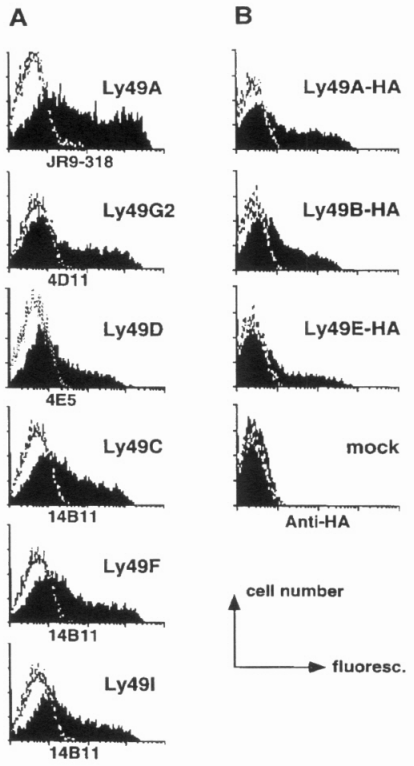


Figure 1. Cell Surface Expression of Ly49 Receptors on COS-7 Transfectants

(A) COS-7 cells were transiently transfected with the indicated Ly49 cDNAs (filled curve) or a negative control (inverted Ly49G2 cDNA, open curve). After 2 days, cells were stained with the indicated FITC-conjugated mAbs.

(B) COS-7 cells transfected with HA-tagged Ly49A, Ly49B, Ly49E, or control cDNA were stained with anti-HA (filled curve) or an isotype control mAb (open curve) and donkey anti-Mlg-PE.

I–deficient blasts from comparison strains that were homozygous for a mutation in the *β2 microglobulin* (*β2m*) gene ($H\text{-}2^{b,\ d,\ or\ k}$) or the *TAP-1* gene ($H\text{-}2^b$) (Figure 2; Table 1; data not shown). In contrast, deficiency of class II molecules had no effect in the assay (in the case of $H\text{-}2^b$ [data not shown]).

In accordance with previous data (Daniels et al., 1994a; Brennan et al., 1996b), Ly49A transfectants bound strongly to Con A blasts of the $H\text{-}2^d$ and $H\text{-}2^k$ but not the $H\text{-}2^b$ haplotype. In eight independent experiments, binding of Ly49A to $H\text{-}2^d$ was consistently higher (about 1.8-fold) than binding to $H\text{-}2^k$ (Figure 2; data not shown). In contrast to a previous report (Brennan et al., 1996b), Ly49A transfectants bound significantly to $H\text{-}2^s$ cells, albeit weakly. Interestingly, Ly49A transfectants also bound Con A blasts of all other haplotypes tested, i.e., $H\text{-}2^{f,q,r,v}$. The results indicate that Ly49A exhibits a broad and often strong MHC reactivity.

Consistent with functional evidence that Ly49G2 interacts with D^d and/or L^d (Mason et al., 1995), Ly49G2-transfected cells reproducibly bound to $H\text{-}2^d$ cells. This

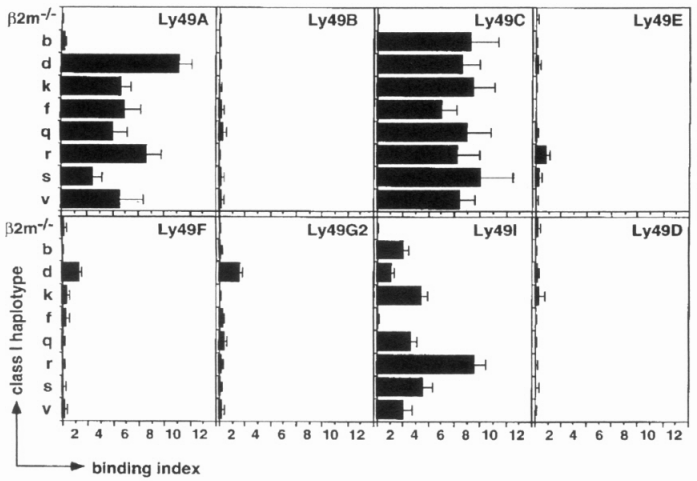


Figure 2. Cell–Cell Adhesion Assay Detects Binding of Ly49 Receptors to Lymphoblasts from MHC Congenic Mice

Ly49-transfected COS-7 cells were adhered to ^{3}H-labeled Con A blasts from B6 *β2m*$^{-/-}$ (β2m$^{-/-}$), B10 (b), B10.D2 (d), B10.BR (k), B10.M (f), B10.Q (q), B10.RIII (r), B10.S (s), and B10.SM (v) mice (MHC haplotype in parentheses). The binding index was determined as "cpm Ly49 transfectants / cpm mock transfectants." Means ± SEM of three to ten independent experiments are shown.

interaction was not detected in a previous binding study (Takei et al., 1997). Binding by Ly49G2 was substantially lower than binding by Ly49A-transfected cells in each of eight independent comparisons. Interestingly, Ly49G2 displayed little if any reactivity toward any of the other seven MHC haplotypes tested, suggesting that Ly49G2 is highly selective for MHC binding.

Ly49C reacted strongly with Con A blasts of all eight MHC congenic strains tested. These data are consistent with previous results showing that Ly49C interacts with K^b and H-$2^{d,k,s}$ molecules expressed on tumor cells (Brennan et al., 1996b) and extend the MHC specificity of Ly49C to new class I alleles. The closely related Ly49I molecule reacted with H-2^r cells to a similar extent as Ly49C. However, Ly49I bound all other MHC haplotypes more weakly than did Ly49C and failed to bind H-2^f cells at all. A previous study failed to detect Ly49I binding to MHC molecules (Takei et al., 1997).

Ly49F, also similar in sequence to Ly49C and Ly49I, bound weakly but consistently to MHC molecules of the H-2^d haplotype but failed to bind cells of all the other haplotypes. Ly49B, Ly49D, and Ly49E displayed little, if any, MHC reactivity to any haplotype in the adhesion assay.

In each case where the binding of Ly49-transfected cells to H-2^d lymphoblasts occurred, the binding could be blocked by corresponding anti-Ly49 mAbs. Thus, binding by Ly49A-transfectants was blocked by mAb A1, binding by Ly49G2 transfectants was blocked by mAb 4D11, and binding by Ly49C, -F, and -I transfectants was blocked by mAb 14B11 (Table 1). Isotype control mAbs had no effect (data not shown).

MHC Locus Specificities of Inhibitory Ly49 Molecules

For more detailed analysis of MHC specificities, Con A blasts from MHC recombinant mice were used in the

Table 1. Class I Specificity of Ly49 Binding in Cell–Cell Adhesion Assay

Adhesion Partners	H-2 KDL[1]	mAb	Transfectants Ly49A	Ly49C	Ly49F	Ly49G2	Ly49I
B6/B10	bb–	-	1.0 ± 0.1	8.0 ± 4.7	1.1 ± 0.1	1.0 ± 0.1	2.3 ± 0.5
B6 β2m$^{-/-}$		-	1.0 ± 0.1	1.1 ± 0.1	1.2 ± 0.1	1.1 ± 0.1	1.0 ± 0.1
B10.D2	ddd	-	7.3 ± 1.4	6.9 ± 1.7	2.2 ± 0.4	2.7 ± 0.2	1.9 ± 0.4
B6-H-2^d-β2m$^{-/-}$		-	1.0 ± 0.1	1.0 ± 0.2	1.6 ± 1.0	1.2 ± 0.1	1.1 ± 0.2
B10.BR	kk–	-	4.2 ± 1.4	8.5 ± 3.5	1.5 ± 0.3	1.0 ± 0.1	3.2 ± 1.1
B10.BR β2m$^{-/-}$		-	0.9 ± 0.2	1.0 ± 0.1	1.4 ± 0.1	1.0 ± 0.1	1.0 ± 0.1
B10.HTG	db–	-	1.5 ± 0.1	3.6 ± 1.4	2.5 ± 0.8	1.3 ± 0.1	2.9 ± 1.0
B10.D2(R107)	bdd	-	9.3 ± 2.3	7.7 ± 2.1	1.0 ± 0.1	3.1 ± 0.5	4.0 ± 1.5
B10.A(2R)	kb–	-	1.0 ± 0.1	7.2 ± 2.7	1.3 ± 0.2	1.1 ± 0.1	1.5 ± 0.1
B10.D2	ddd	Ly49[2]	1.1 ± 0.1	1.3 ± 0.1	1.1 ± 0.2	1.3 ± 0.2	0.9 ± 0.1
B10.D2	ddd	D^d [3]	1.2 ± 0.4	5.4 ± 2.0	1.1 ± 0.4	0.8 ± 0.0	0.9 ± 0.1
B10.BR	kk–	D^k,K^k [4]	1.1 ± 0.2	1.2 ± 0.2	0.8 ± 0.3	0.7 ± 0.1	0.6 ± 0.2
B6/B10	bb–	K^b [5]	nd	0.8 ± 0.2	0.9 ± 0.1	nd	0.8 ± 0.1

Data represents mean binding index ± SEM of three to six independent semiquantitative adhesion experiments with the indicated COS-7 transfectants and Con A blasts from the listed strains. The index was calculated in reference to COS-7 cells transfected with an inverted Ly49G2 cDNA.

[1] The H-2 designations refer to allelic variants at K, D, and L. The hyphen means that no L product is made in this strain.

[2] Anti-Ly49 mAbs used for blocking (50 μg/ml) were A1 (anti-Ly49A), 14B11 (anti-Ly49C,F,I), and 4D11 (anti-Ly49G2). Anti–class I mAbs were 34-5-8S[3] (anti-D^d α1/α2), 15-1-5P[4] (anti-K^k, D^k), and Y3[5] (anti-K^b). nd, not determined.

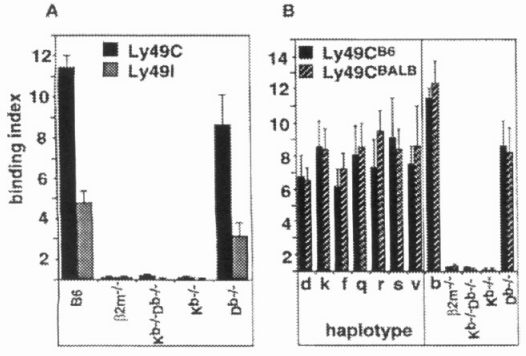


Figure 3. MHC Specificities of Ly49I, Ly49C^{B6}, and Ly49C^{BALB}

(A) Ly49C and Ly49I bind to K^b molecules on lymphoblasts but not appreciably to D^b molecules. The adhesion assay was performed with Ly49C or Ly49I transfectants and Con A blasts from B6 mice or B6 mice deficient for β2m, D^b, K^b, or both K^b and D^b. Means ± SEM of four to six determinations from two to four independent experiments are shown.

(B) Ly46C^{B6} and Ly49C^{BALB} are indistinguishable in MHC binding. The adhesion assay was performed as described in Figure 2. Means ± SEM of three to six independent experiments are shown.

adhesion assay, and anti-class I mAbs were employed to block binding (Table 1). Confirming the specificity of Ly49A for D^d was the finding that anti-D^d mAb blocked binding of Ly49A transfectants to B10.D2 cells (Table 1). Furthermore, Ly49A-transfected cells bound to B10.D2(R107) cells, which express K^b, D^d, and L^d, and did not bind to B10.HTG cells, which express K^d but are otherwise H-2^b (Table 1). Binding of Ly49A transfectants to H-2^k cells was blocked with a mAb that reacts with both K^k and D^k. Significantly, the Ly49A transfectants did not bind to B10.A(2R) cells (K^k/D^b), consistent with previous genetic results indicating that Ly49A interacts with D^k and not K^k (Karlhofer et al., 1994).

As was the case for Ly49A, Ly49G2 transfectants bound to B10.D2(R107) cells but not to B10.HTG cells, implicating D^d and/or L^d in binding. Binding to H-2^d cells was blocked to background levels by mAb specific for $D^d\alpha1/\alpha2$ (Table 1). These data confirm that D^d is a ligand for Ly49G2, in contrast to a report that questioned this assignment (Takei et al., 1997). The complete blockade with anti-D^d mAb suggests that L^d is not an effective ligand for Ly49G2, in contrast to a previous study suggesting that both molecules are ligands (Mason et al., 1995). Since Ly49G2 bound significantly only to D^d among eight haplotypes tested, this receptor exhibits a very high degree of selectivity in MHC binding.

The binding of Ly49C and Ly49I transfectants to H-2^b Con A blasts was completely blocked with a specific anti-K^b mAb, confirming that K^b is essential for both of these interactions (Table 1; Brennan et al., 1996b). This conclusion was strongly supported by analysis of binding of Con A blasts from mice with targeted mutations in K^b, D^b, or both (Perarnau et al., 1999). Cells with the K^b mutation failed to bind to either Ly49C- or Ly49I-transfected cells (Figure 3A). In contrast, cells with the D^b mutation bound nearly as well as wild-type cells. Therefore, K^b is the major H-2^b class I ligand for Ly49C and Ly49I and D^b plays no role or a very minor one.

Binding of Ly49C or Ly49I to H-2^k cells was completely blocked by an antibody that reacts with both K^k and D^k, demonstrating interactions with one or both of these class Ia molecules. Binding of Ly49C to H-2^d blasts was only partially blocked by anti-D^d mAb, consistent with the previous report that K^d may also bind Ly49C (Brennan et al., 1996b). Binding of Ly49I to H-2^d molecules was reduced to background with anti-D^d mAb, arguing that D^d is the major H-2^d ligand for Ly49I. However, since Ly49I also bound significantly to B10.HTG blasts (K^dD^b), K^d and/or D^b could not be ruled out as ligands for Ly49I. H-2^d was the only haplotype displaying reactivity with Ly49F above background. Since anti-D^d mAb blocked binding to B10.D2 cells, but Ly49F also reacted with B10.HTG (K^dD^b) cells, the locus specificity of Ly49F remains uncertain.

Different Ly49C Alleles Bind the Same Class I Molecules

The Ly49C protein exhibits allelic diversity, differing in four amino acids between the B6 and BALB strains (Brennan et al., 1996a). We compared binding of Ly46C^{B6} and Ly49C^{BALB} transfectants in the adhesion assay. As shown in Figure 3B, we observed no difference in class I reactivity between Ly46C^{B6} and Ly49C^{BALB}. Supporting this finding, Ly46C^{B6} and Ly49C^{BALB} transfectants were also bound equally well by soluble MHC class I/peptide tetramers of the H-$2^{b,d,k}$ haplotypes (data not shown). Therefore, we find no evidence for differences in the specificity of these allelic Ly49C receptors.

Avidity for Class I Determines the Inhibitory Function of Ly49A and Ly49G2

To address whether the differences in class I binding by Ly49 molecules as detected in our assay is related to the degree of inhibitory function, we correlated binding in the cell–cell adhesion assay with the capacity of MHC-different spleen cells to inhibit the function of Ly49-expressing lymphocytes. For these experiments, we took advantage of the fact that transgenically expressed Ly49 molecules on T cells readily inhibit the proliferative response of the T cells to allogeneic stimulator cells expressing cognate class I ligands (Held et al., 1996a). The use of T cells also has the important advantage that very few of these cells express additional endogenously encoded Ly49 receptors that might confound the specificity analysis. We therefore employed Ly49A transgenic mice (Held et al., 1996a) and a new Ly49G2 transgenic line as a source of T cells to examine MHC specificity. Similar to the Ly49A transgene, the Ly49G2 transgene was expressed on all T cells at homogeneously high levels (Figure 4A). T cells from transgenic or nontransgenic littermates on the B6 (H-2^b) background were used as responder cells with spleen cells from MHC congenic mice as allogeneic stimulator cells.

The relative inhibition of T cell proliferation by the Ly49A transgene corresponded quite well with the class I avidity of Ly49A as determined by the adhesion assay (Figure 4). Inhibition was most prominent with H-2^d stimulator cells and least prominent with H-2^s stimulator cells. The other haplotypes showed intermediate inhibition effects that correlated with their binding properties, except in the case of the H-2^f stimulator cells, which

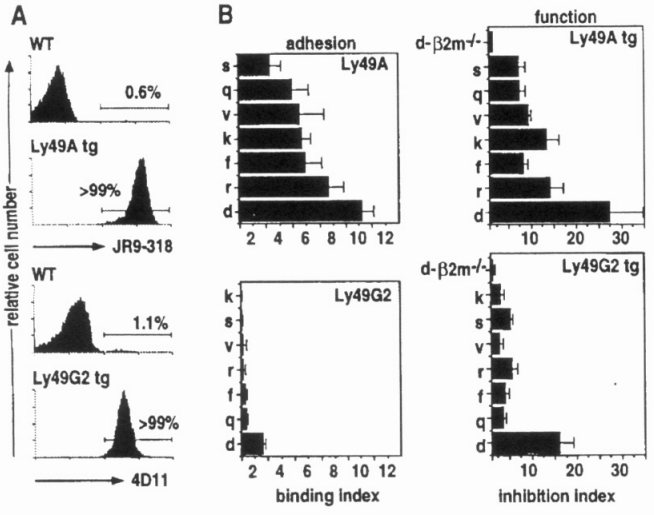


Figure 4. The Inhibitory Function of Ly49A and Ly49G2 Is Related to the Extent of Binding as Determined in the Cell–Cell Adhesion Assay

(A) Ly49A and Ly49G2 transgenic mice express the corresponding Ly49 receptor at high levels on T cells. JR9–318 (upper panel) or 4D11 (lower panel) staining of gated $CD3^+$ nylon wool–passed lymph node cells is shown for Ly49A transgenic mice (Ly49A tg), Ly49G2 transgenic mice (Ly49G2 tg), or nontransgenic littermates (wt).

(B) Relationship of class I binding and inhibitory function. The T cells represented in (A) were used as responder cells in mixed lymphocyte reactions with irradiated spleen cells from the indicated MHC congenic strains as stimulator cells. Relative binding data are taken from Figure 2. The inhibition index was determined as the stimulation index with nontransgenic responder cells / the stimulation index with transgenic responder cells. The data represent means ± SEM of four independent experiments, each determined in triplicate, except in the case of the H-2^d $\beta 2m^{-/-}$ stimulator cells, where the data represent the means of two (Ly49A tg) or three (Ly49G2 tg) experiments.

exhibited slightly less than expected levels of inhibition. For Ly49G2 transgenic responder cells, the T cell response was most strongly inhibited by H-2^d stimulator cells, in agreement with D^d being the dominant ligand for Ly49G2 in the adhesion assay. The inhibition was relatively strong in comparison to the relatively weak binding of Ly49G2 to H-2^d cells in the adhesion assay (see Discussion). The other haplotypes exhibited only weak inhibitory function. The extent of inhibition observed with H-2^d-$\beta 2m^{-/-}$ stimulator cells, which present foreign MHC class II molecules but no potentially inhibitory class I molecules, was minimal in the case of both Ly49A and Ly49G2 transgenic responder cells. Thus, the receptors do not cause significant inhibition in the absence of class I ligands. Taken together, these results demonstrate a clear relationship between the class I molecules that bind in the cell–cell adhesion assay and those that cause inhibition. The results with Ly49A allow the additional conclusion that inhibitory signaling is not an all or none phenomenon, since even weakly binding MHC haplotypes (e.g., H-2^s) exhibited some inhibitory function. Therefore, the other instances of relatively weak binding of inhibitory Ly49 receptors to class I molecules are likely to have functional significance.

Binding of Ly49 Molecules to Soluble MHC Class I/Peptide Tetramers

To complement the cell–cell adhesion assay in addressing the MHC specificities of Ly49 receptors, we examined the binding of fluorochrome-labeled tetrameric class I/β2m/peptide complexes to Ly49-transfected cells. This assay represents a more direct measurement of binding to Ly49 receptors, free from the hypothetical influence of non-MHC molecules on class *I–Ly49 interactions*, and also allows an assessment of the role of glycans and MHC-bound peptides in Ly49 binding.

Ly49-transfected COS-7 cells were stained with K^b, D^b, K^d, D^d, L^d, and D^k tetramers and analyzed by flow cytometry (Figure 5A). All of the tetramers except L^d exhibited specific binding to transfected COS cells expressing at least one of the inhibitory Ly49 receptors. No binding of Ly49-transfected COS-7 cells to nonclassical Qa-1^b tetramers, the ligand for CD94/NKG2A (Vance et al., 1998), was observed (data not shown). These data demonstrate direct binding of class Ia molecules to Ly49 receptors. Furthermore, since the class I molecules (with the exception of D^d) were generated in bacteria and are not glycosylated, the results indicate that N-glycosylation of class I molecules is not required for Ly49 binding.

As expected, Ly49A transfectants stained specifically with D^d and D^k tetramers but not with the other six tetramers. The binding of D^k tetramers corroborates the genetic evidence for this interaction (Karlhofer et al., 1994). Ly49G2 transfectants displayed weak but selective and reproducible reactivity with D^d tetramers. Binding of the D^d tetramers to Ly49A or Ly49G2 was blocked by the anti-Ly49A mAb JR9-318 or the anti-Ly49G2 mAb 4D11, respectively (Figure 5B). The Ly49G2 transfectants failed to stain with the K^b, D^b, K^d, and D^k tetramers. No reactivity with L^d tetramers was observed in two experiments (e.g., Figure 5), but very weak reactivity was observed in two others (data not shown). The L^d tetramers effectively stained LCMV-immune T cells, demonstrating their functionality (data not shown). In line with the cell–cell adhesion results, these data imply that L^d does not bind appreciably to Ly49G2.

Interestingly, Ly49C reacted with all class I tetramers except L^d, corroborating the promiscuity of this receptor observed in the cell–cell adhesion assay. Binding to K^b, K^d, D^d, and D^k is in accordance with the data from the adhesion assay and with results published previously (Brennan et al., 1996b). However, binding to D^b was

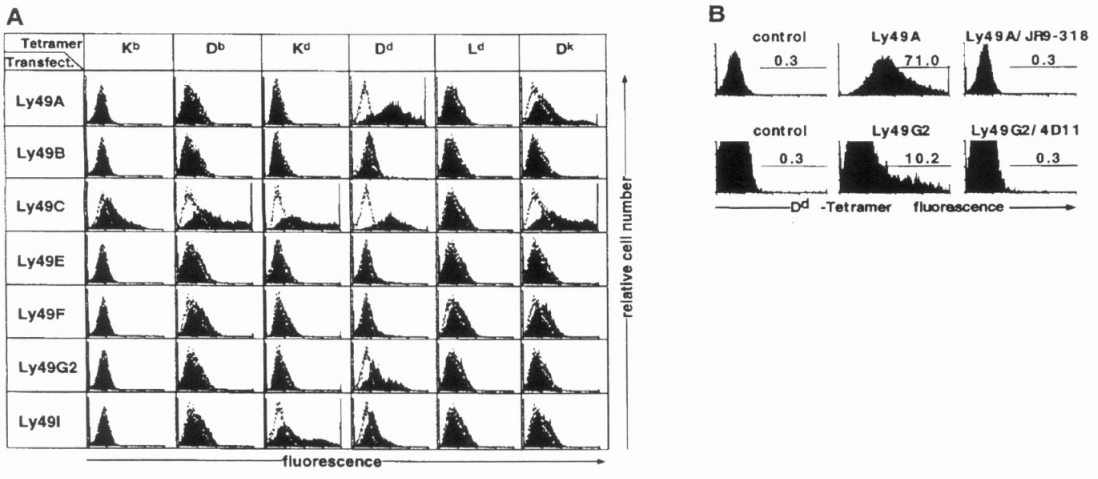


Figure 5. Distinct Ly49 Molecules Bind to MHC Class I Tetramers of the H-2^b, H-2^d, and H-2^k Haplotypes

(A) Ly49-transfected COS-7 cells were stained with MHC class I tetramers complexed to β2m and the following peptides: K^b, SEV9 (Sendai virus); D^b, gp33 (LCMV); K^d, JAK1 (self); D^d, gp49 (HIV); L^d, NP118 (LCMV); D^k, MT389 (polyoma virus). The filled curves represent Ly49 transfectants and the open curves control (inverted Ly49G2) transfectants. Quantitative comparisons are best made within a column, each of which represents an independent experiment. Ly49 expression on transfectants varied somewhat in different experiments. Histograms are representative of two to four experiments for each group.

(B) Anti-Ly49 mAbs inhibit staining with D^d tetramers. Ly49A-transfected or Ly49G2-transfected COS-7 cells were stained with D^d tetramers in the presence or absence of anti-Ly49A (JR9–318) or anti-Ly49G2 (4D11) mAbs, respectively. Staining of control Ly49G2^{inverse} transfectants is also shown.

surprising since Ly49C bound poorly, if at all, to D^b molecules expressed on Con A blasts in the cell–cell adhesion assay (Table 1; Figure 3) (Brennan et al., 1996b). The capacity of Ly49C transfectants to bind D^b tetramers was probably not due to unusual properties of the specific peptide bound to D^b, because equivalent binding was observed with D^b tetramer preparations containing two different bound peptides (LCMV NP396 and LCMV gp33 [data not shown]). D^b tetramers did not stain other Ly49 transfectants (Figure 5), and D^b tetramer binding to Ly49C was blocked by D^b-specific mAb B22–249 but not by K^b-specific mAb Y3 (data not shown). We conclude that Ly49C binds well to tetrameric recombinant D^b/peptide complexes but poorly, if at all, to D^b on lymphoblasts (see Discussion).

In two instances, interactions documented in the cell–cell adhesion assay were not detected in the tetramer studies. Ly49I-transfectants were stained well by K^d tetramers, weakly by D^d tetramers, and at background levels by K^b tetramers. The absence of reactivity with K^b tetramers contrasts with the unequivocal role of K^b in the adhesion assay, in which Ly49I transfectants failed to bind K^b knockout cells, and binding to wild-type cells was blocked with anti-K^b mAb (Table 1; Figure 3) (see Discussion). Ly49F, despite being a weak H-2^d-specific receptor in the adhesion assay (Figure 2; Table 1), was not detected with the K^d, D^d or L^d tetramers. In agreement with the adhesion experiments, Ly49B and Ly49E transfectants did not react significantly with any of the class I tetramers.

Ly49I Binding Depends on the Peptide Bound to K^d

We were able to examine the role of peptide in the binding of Ly49C and Ly49I to K^d with the use of a panel of K^d-peptide tetramers containing five unrelated peptides (Figure 6). Ly49C-transfected COS-7 cells bound specifically and equivalently to soluble K^d tetramers complexed to any of five different peptides. These data indicate that Ly49C exhibits little, if any, peptide specificity and further demonstrates the integrity of all

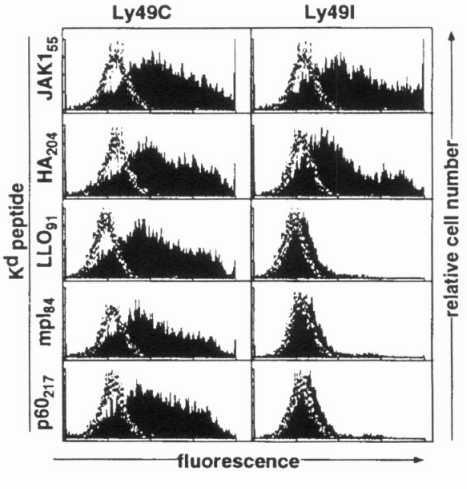


Figure 6. Ly49I but Not Ly49C Discriminates Peptides Bound to K^d

Ly49C, Ly49I (filled curves), or control (open curves) transfected COS-7 cells were stained with PE-conjugated K^d/β2m tetramers complexed to the indicated peptides. Identical results were obtained in a repeat experiment.

the tetramers. Interestingly, however, only two of the five peptides allowed binding of K^d to Ly49I. These data demonstrate that Ly49I can discriminate peptides presented by K^d.

Discussion

Evidence for a physical interaction with MHC class I molecules has previously only been reported for Ly49A with D^d and H-2^k and for Ly49C with K^b, K^d, D^d, and H-2^s gene products (Daniels et al., 1994a; Kane, 1994; Brennan et al., 1996b). MHC interactions with Ly49G2 and Ly49I have been inferred from functional studies (Mason et al., 1995; Yu et al., 1996) but have been questioned based on binding studies (Takei et al., 1997). Our cell–cell adhesion studies demonstrate several binding interactions that were not detected in previous binding studies, including Ly49G2 to H-2^d, Ly49A to H-2^s, and Ly49I to H-$2^{b,d,k,s}$. Most of these interactions were relatively weak, perhaps explaining the earlier failure to detect them. The Ly49G2–H-2^d interaction can be attributed mainly to the D^d class I molecule, since binding was blocked with anti-D^d mAb, and D^d but not L^d tetramers bound reproducibly to Ly49G2-transfected cells. The apparent lack of a role of L^d is at odds with some earlier functional results (Mason et al., 1995) but consistent with the results of a recent transgenic study implicating D^d but not L^d as a ligand inducing "missing self" reactivity in Ly49G2-expressing NK cells in vivo (Johansson et al., 1998). The cell–cell adhesion data also show that Ly49A and Ly49C bind strongly to a variety of additional MHC class I alleles, suggesting that common structural motifs define the class I specificities of these molecules (Figures 2–4; Table 1). In addition, evidence from the binding assay (Figure 3B) argues that allelic differences between Ly49C receptors do not contribute significantly to differences in class I specificity. More importantly, we report instances of MHC binding by Ly49I and Ly49F (Figures 2–4; Table 1).

It is interesting that some receptors exhibit such broad MHC cross-reactivity (e.g., Ly49C and Ly49A) while others are quite selective in MHC binding (e.g., Ly49G2). The more cross-reactive receptors provide an explanation for how such a small number of receptors can cover *the much larger* number of MHC alleles present in the population. The more selective receptors may be advantageous, since they would allow a host to detect the selective loss of particular MHC alleles. The receptors that failed to bind well to any of the tested MHC haplotypes, Ly49B and Ly49E, may have MHC ligands that were not represented in our panel. Ly49E in particular is very similar in sequence to other MHC-reactive Ly49 receptors (Ly49C, Ly49I, and Ly49F), suggesting a role in MHC recognition. Indeed, we observed very weak binding of Ly49E to H-2^r molecules in some experiments (Figure 2; data not shown). Ly49B, on the other hand, is unique in displaying by far the least homology to the other Ly49 family members (Wong et al., 1991), raising the possibility that it has a distinct function or specificity.

In the cases of Ly49A and Ly49G2, we provide evidence that the extent of binding observed in the cell–cell adhesion assay correlates with the extent of functional inhibition (Figure 4). By performing the assays with T cells, most of which do not express endogenously encoded Ly49 receptors (Ortaldo et al., 1998; our unpublished data), the contributions of other Ly49 receptors to inhibition is largely avoided. The most significant aspect of these data is that even weak binding by Ly49A can lead to inhibitory effects, suggesting that inhibition is not an all or none threshold phenomenon, but rather varies continuously with receptor avidity. The functional cross-reactivity of Ly49A with multiple MHC haplotypes was confirmed by limited analysis demonstrating that IL-2-activated NK cells expressing transgenic Ly49A are inhibited by Con A blasts from mice of several different MHC haplotypes (data not shown). The weak binding of Ly49G2 to D^d also resulted in significant functional inhibition in the T cell assay. The inhibition was relatively strong considering that binding to H-2^d cells by Ly49G2 was weaker than binding by Ly49A. This may occur because transgene-directed Ly49G2 levels were elevated on T cells compared to their normal expression level on NK cells (data not shown), or because Ly49G2 signaling is more efficient than Ly49A signaling. The finding that weak binding events lead to detectable inhibitory effects argues against the possibility that the other binding events we document are in some cases of too low an avidity to be functionally relevant. Equally pertinent, results from comparing NK cells that express Ly49A or Ly49G2 alone or together suggest that inhibitory effects mediated by two or more receptors on the same NK cell are cumulative (T. H. and D. H. R., unpublished data). Since NK cells commonly coexpress multiple Ly49 receptors, even the weak binding interactions we observe in the case of individual Ly49 receptors are likely to combine to produce a stronger inhibitory effect. Therefore, although we have not documented functional effects for all the binding events observed in the cell–cell adhesion assay, they probably play physiologically significant roles in the overall inhibitory interaction.

The binding data provide potential explanations for two previous observations. Of the seven well-characterized Ly49 receptors with ITIMs, only Ly49C and Ly49I bind significantly to H-2^b class I molecules. In contrast, five of the seven inhibitory Ly49 receptors react appreciably with H-2^d class I molecules. The large number of H-2^d-specific inhibitory receptors and the relative dearth of H-2^b-specific receptors may provide an explanation for the fact that NK cells in H-$2^{b/d}$ mice reject H-2^d bone marrow very inefficiently, while they reject H-2^b bone marrow grafts readily (Murphy et al., 1990). Furthermore, the abundance of H-2^d-specific receptors could account for the counterintuitive observation that the frequencies of NK cells that express Ly49A or Ly49G2 are marginally lower in H-2^d mice than in H-2^b mice (Held et al., 1996b), even though it might be expected that cells expressing H-2^d-specific (i.e., self-specific) receptors, as a group, would be more frequent in H-2^d mice. Mathematical modeling of proposed NK cell "education" processes led to the prediction that the observed decrease in the frequency of cells expressing any one H-2^d-specific receptor in H-2^d mice would result under two conditions (Vance and Raulet, 1998). First, there must exist mechanisms to disfavor the coexpression of multiple self–class I–specific receptors by NK cells. This premise was supported by the finding that expression of a Ly49A transgene in all developing NK cells of H-2^d mice clearly

inhibits the expression of a second H-2^d-specific receptor, Ly49G2 (Held and Raulet, 1997). Second, H-2^d-specific receptor genes must be significantly more common than H-2^b-specific receptor genes in the germline receptor set (Vance and Raulet, 1998). The present data support the latter prediction.

The demonstration that class Ia tetramers bind to Ly49 receptors indicates that caution must be applied when using these reagents to detect specific T cell receptors, since the reagents may instead be detecting Ly49 receptors expressed on T cells (Ortaldo et al., 1998). The tetramer results have several implications concerning the nature of the Ly49–class I interaction. First, they go farther than previous studies in providing direct evidence for a physical interaction between class I molecules and Ly49 receptors. Second, since all tetramers except D^d were produced in bacteria and are therefore unglycosylated, the results provide compelling evidence that Ly49 binding to most class I molecules does not require glycosylation of the class I molecules. That class I–linked glycans might play a role in Ly49 binding had been suggested by the lectin-like structure of these receptors. Our results are in accord with previous studies demonstrating that mutagenesis of the acceptor sites for N-glycosylation in D^d did not prevent Ly49A binding (Lian et al., 1998; Matsumoto et al., 1998). The tetramer data do not, however, rule out the possibility that interaction of Ly49 receptors with class I-linked glycans might alter the affinity or specificity of the interaction.

The third significance of the tetramer binding studies concerns the role of MHC-bound peptides in the interaction. Previous results demonstrated that Ly49A does not discriminate D^d-bound peptides as tested in functional studies (Correa and Raulet, 1995; Orihuela et al., 1996). Along the same lines, the new data presented here indicate that Ly49C exhibits little, if any, capacity to discriminate peptides bound to K^d (Figure 6) or D^b (Figure 5; data not shown). Ly49I, however, reacted with only two of the five K^d–peptide complexes tested. The three complexes that did not bind to Ly49I bound well to Ly49C, demonstrating their integrity. These complexes showed essentially no reactivity with Ly49I-transfected cells, suggesting an appreciable reduction in the affinity compared to Ly49I binding to the other two K^d/peptide tetramers. The peptides that were compatible with Ly49 binding included a viral peptide (influenza HA^{204}) and a self-peptide ($JAK1^{55}$), indicating no general theme in the origin of active peptides. It remains to be determined whether the discrimination of peptides by Ly49I observed with tetramers also holds in functional tests and whether the peptide residues actually provide binding energy to the interaction or, alternatively, sterically impede the interaction.

A surprising finding was that D^b-tetramers efficiently stained Ly49C-transfected cells (Figure 5A). No appreciable binding of Ly49C to D^b could be detected in the cell–cell adhesion assay by antibody blocking or by employing cells from $K^{b-/-}$ mice that express D^b normally (Figures 3 and 4). Furthermore, previous functional (Yu et al., 1996) and binding (Brennan et al., 1996b) studies *have failed to detect an* interaction between Ly49C and D^b. D^b tetramers containing two distinct peptides stained Ly49C transfectants, arguing against the possibility that the failure to bind D^b on cells reflects a requirement for a specific peptide bound to D^b (Figure 5A; data not shown). The possibility must therefore be considered that the tetrameric D^b molecules produced in bacteria can in some cases exhibit different binding properties than D^b expressed in mammalian cells. One possibility is that the absence of glycans on the bacterially produced D^b reveals an epitope for Ly49C. Thus far, however, we have been unable to enhance binding of Ly49C transfectants to $K^{b-/-}$ cells by treating the latter cells with glycosidase (data not shown). It is also conceivable that the use of human β2m instead of mouse β2m in these tetramers could account for the unexpected binding behavior.

Another apparently anomalous finding in the tetramer studies was that D^d and K^b tetramers failed to bind well to Ly49I despite the clear evidence of binding observed in the cell–cell adhesion assay. Possibly relevant to these data is our observation that binding of D^d tetramers to $Ly49A^+$ NK cells appears to be highly sensitive to Ly49A surface expression levels. By mAb staining, the levels of Ly49A are reduced by 2- to 3-fold on NK cells from H-2^d mice compared to those from H-2^b mice. However, while the D^d tetramers are very effective in staining $Ly49A^+$ NK cells from H-2^b mice, they barely stain $Ly49A^+$ NK cells from congenic H-2^d mice (data not shown). Thus, tetramer binding appears to fall off sharply at a threshold Ly49 receptor density, consistent with the multivalent binding requirements expected of tetramers. Variations in affinity and in the levels of different Ly49 receptors on transfected cells could therefore explain why some class I–Ly49 interactions cannot be detected in our tetramer assays. Another possibility is that the peptides in the grooves of the D^d and K^b tetramers are not compatible with Ly49I binding. This explanation is plausible in light of the finding that Ly49I discriminates K^d tetramers complexed with different peptides.

An additional conclusion of our study was that the Ly49C proteins encoded by B6 and BALB strains, which differ in four amino acids, are indistinguishable in MHC specificity in the cell adhesion assay. Supporting this finding, $Ly46C^{B6}$ and $Ly49C^{BALB}$ also reacted equally with soluble MHC class I/peptide tetramers of the H-$2^{b,d,k}$ haplotypes (data not shown). Based on functional studies, the possibility had been raised that these proteins differ in specificity (Yu et al., 1996), but it now appears that there must be another explanation for the functional results (George et al., 1997).

Finally, no interaction of Ly49D with D^d-expressing cells was detectable in our cell–cell adhesion studies, despite the functional evidence for this interaction. Furthermore, the D^d tetramers failed to detectably stain Ly49D-transfected COS-7 cells (data not shown). The data raise the possibility that the D^d–Ly49D interaction is weaker than some of the inhibitory interactions studied here. However, the possibility remains that the poor binding reflects a somewhat lower level of Ly49D on transfected cells than the levels of the inhibitory receptors, perhaps due to the requirement for association with cotransfected DAP12, or that Ly49D is more peptide or glycan specific than the inhibitory receptors. Ly49H-transfectants also failed to bind to class I molecules in either the cell–cell adhesion assay or the tetramer studies (data not shown). However, the significance of these

data is suspect, because the level of Ly49H on transfectants was substantially lower than the levels of the other Ly49 receptors studied (data not shown).

These results provide the most comprehensive picture to date of MHC binding by different Ly49 receptors and go a long way toward providing a basis for analyzing the overall repertoire of MHC specificities expressed by NK cells in a given mouse strain. A complete picture of the repertoire will require the incorporation of recently discovered inhibitory receptors, including the CD94/NKG2 receptors that react with a peptide derived from signal sequences of D region molecules that is presented by the Qa-1^b class Ib molecule (Vance et al., 1998). In addition, additional Ly49 genes have been discovered at the genomic level in B6 mice (McQueen et al., 1998); whether these receptors are functionally expressed remains to be determined. Once a complete panel is available, several other important questions can be addressed, such as whether NK cell self-tolerance can be accounted for by the expression of self-MHC-specific inhibitory receptors on each NK cell and whether congenic mice of different MHC haplotypes exhibit distinct inhibitory receptor repertoires.

Experimental Procedures

Animals and Generation of Transgenic Mice

C57BL/6J (H-2^b, B6), C57BL/10J (H-2^b, B10), B10.D2/nSnJ (H-2^d, B10.D2), B10.BR/SgSnJ (H-2^k, B10.BR), B10.M/Sn (H-2^f, B10.M), B10.Q/SgJ (H-2^q, B10.Q), B10.RIII(71NS)/Sn (H-2^r, B10.RIII), B10.S/SgMcdJ (H-2^s, B10.S), B10.SM(70NS)/Sn (H-2^v, B10.SM), B10.HTG/2Cy (H-2^g, B10.HTG), B10.D2(R107)/Eg (H-2^{r7}, B10.D2(R107)) and B10.A(2R)/SgSnJ (H-2^{h2}, B10.A(2R)) mice were purchased from the Jackson Laboratory. B6-$\beta 2m^{-/-}$ mice (Zijlstra et al., 1990) had been backcrossed five times to B6 mice before intercrossing. B10.BR-$\beta 2m^{-/-}$ and B10.D2-$\beta 2m^{-/-}$ mice were previously described (Held et al., 1996b), as were Ly49A transgenic mice (line 2) (Held et al., 1996a). $K^{b-/-}$, $D^{b-/-}$, and $K^{b-/-}D^{b-/-}$ mice of the B6 background have been described (Perarnau et al., 1999). For Ly49G2 transgenic mice, the complete coding sequence of the Ly49G2 cDNA (identical to the sequence published by Smith et al., 1994) was subcloned into the pHSE expression cassette (Pircher et al., 1989) and injected into fertilized (B6 × CBA/J F2) eggs. The line described here (line 5) was backcrossed at least four times to B6 mice and contained approximately eight transgene copies as determined by Southern blotting.

Cloning of Ly49 cDNAs

$Ly49^{B6}$ cDNAs were cloned by RT–PCR from total B6 LAK cell or EL-4 RNA using Ly49 family member specific primers. Ly49A (amino acid sequence identical to that in Held et al., 1995), Ly49B (sequence from Wong et al., 1991), Ly49C (sequence from Stoneman et al., 1995), Ly49E, Ly49F and Ly49G2 (sequences from Smith et al., 1994), and Ly49I (identical to sequence by Brennan et al., 1996a) cDNAs were obtained. An HA-peptide epitope tag (amino acid sequence YPYDVPDYA) was added to the C termini of Ly49B, Ly49E, and Ly49A (the latter only for staining comparisons in Figure 1). The $Ly49C^{BALB}$ cDNA (sequence from Wong et al., 1991) was cloned from a C.B-17 SCID LAK cell cDNA library (kindly provided by P. Mathew and V. Kumar). The cDNAs were subcloned into the pME18S expression vector.

Adhesion Assay

At day −1, COS-7 cells (CRL1651, purchased from ATCC) were seeded at 2 × 10^6 cells/T75 flask. At day 0 (50%–80% confluency), they were transfected with 5 μg of pME-Ly49 (cDNA) or pME-$Ly49G2^{inverse}$ as a control using lipofectamine (GIBCO–BRL) according to the manufacturer's instructions. At day 1, the transfectants were passaged into 6-well plates (Falcon). At day 0, RBC-depleted spleen cell suspensions were cultured at 2 × 10^6/ml with 2.5 μg/ml concanavalin A (Con A). At day 1, they were pulsed with 2.5 μCi/ml ^{3}H-thymidine (Amersham). At day 2, washed COS-7 transfectants were overlaid in 1.5 ml medium with Con A blasts at 1–5 × 10^6 cells/ml. After 2 hr incubation at 37°C, the wells were washed gently 5 times with prewarmed RPMI medium to remove unbound cells. The remaining conjugates were lifted from the plates using PBS/trypsin and transferred into 96-well plates. The plates were harvested automatically and the radioactivity was determined using a β-counter. Typically, approximately 1%–2% of the input radioactivity bound to mock transfectants; in the case of a strong binding interaction, 10%–20% of the input radioactivity bound to the corresponding Ly49-transfected COS-7 cells. The binding index was determined as cpm Ly49-cDNA transfected well/cpm $Ly49G2^{inverse}$ transfected well. COS-7 transfectants were monitored for Ly49 surface expression by staining with anti-Ly49 mAbs after lifting the cells with PBS/0.02%EDTA.

MHC Class I Tetramers

The following tetramers were generated as described in bacteria in association with human β2m and were conjugated to PE or APC (Altman et al., 1996; Busch et al., 1998): K^b, SEV9 (Sendai virus); D^b, gp33 (LCMV); D^b, NP396 (LCMV); K^d, JAK1 55–63 (mouse); K^d, HA 204–212 (influenza virus); K^d, LLO 91–99 (*Listeria*); K^d, p60 217–225 (*Listeria*); K^d, mpl 84–92 (*Listeria*); L^d, NP118 (LCMV); D^k, MT389 (polyoma virus).

D^d tetramers were produced in insect cells, in association with mouse β2m. The murine β2m and D^d cDNAs were cloned into the dual expression transfer vector pAcUW31. The sequences encoding the leader and extracellular domains of D^d (ending at W274) were followed by sequences encoding a His tag and a BirA biotinylation sequence (LNDIFEAQKIEWH). Recombinant D^d was produced using a baculovirus expression system (Summers and Smith, 1987). pHIV peptide (RGPGRAFVTI, 0.25 mg/ml) was added to the supernatants of infected High 5 cells (Invitrogen) and stirred slowly for 12 hr at 4°C. Recombinant D^d was purified by Ni-NTA superflow chromatography (Qiagen) followed by gel filtration chromatography using a Superdex 200 FPLC column (Pharmacia). The integrity and specificity of all tetramers was confirmed by staining antigen-specific T cells or T cell lines.

Antibodies

The following mAbs were used: 4D11 (anti-Ly49G2 [Mason et al., 1995]), JR9–318 (anti-Ly49A [kindly provided by J. Roland]), A1 (anti-Ly49A [Nagasawa et al., 1987]), 4E5 (anti-Ly49D [Mason et al., 1996]), 14B11 (anti-Ly49C, Ly49F, Ly49H, Ly49I [Corral et al., 1999]), 34-5–8S (anti-$D^d\alpha 1/\alpha 2$ [ATCC HB-102]), 15–1–5P (anti-D^k, K^k [ATCC HB-53]), and Y3 (anti-K^b [ATCC HB-176]). Purified mAbs were conjugated to biotin or FITC. Anti-CD3-PE was purchased from Pharmingen, anti-HA from BAbCO, and donkey-anti-mouse Ig-PE (DaMIg-PE) from Jackson.

Staining and Flow Cytometry

Cell staining with antibodies was performed as described (Held et al., 1996b). For tetramer staining, the cells were preincubated for 20 min with 500 μg/ml streptavidin (Molecular Probes), washed, and stained for 1 hr with PE- or APC-conjugated tetramers. They were washed twice and fixed in PBS/1% formaldehyde before analysis.

Mixed Lymphocyte Reaction

Nylon wool–passed lymph node cells (>85% $CD3^+$) from H-2^b transgenic and nontransgenic mice were used as responder cells at 1 × 10^5 cells/round-bottom well with 5 × 10^5 irradiated (2500 rad) spleen cells as stimulators. Cultures were pulsed with 0.5 μCi ^{3}H-thymidine on day 3 and harvested on day 4. Stimulation indices (SI) were determined as cpm with allogeneic stimulator cell / cpm with H-2^b stimulator cells. From these values we determined the "inhibition index" as SI for nontransgenic responder cells / SI with transgenic responder cells.

Acknowledgments

We thank Peter Snow for production of the recombinant D^d from baculovirus, Vinay Kumar and Porunelloor Mathew for the cDNA

library, Werner Held for Ly49A transgenic mice, Chern-sing Goh and Ann Lazar for technical assistance, and Russell Vance for critical reading of the manuscript. T. H. was a recipient of a Research Fellowship from the Deutsche Forschungsgemeinschaft, C. W. M. is the recipient of a Cancer Research Institute/Chase Manhattan Bank fellowship, and D. H. B. is a recipient of a Research Fellowship from the Howard Hughes Medical Institute. F. A. L. was supported by the Institut Pasteur. This work was supported by National Institutes of Health grant RO-1AI35021 to D. H. R.

Received March 8, 1999; revised May 20, 1999.

References

Altman, J.D., Moss, P.A.H., Goulder, P.J.R., Barouch, D.H., McHeyzer-Williams, M.G., Bell, J.I., McMichael, A.J., and Davis, M.M. (1996). Phenotypic analysis of antigen-specific T lymphocytes. Science *274*, 94–96.

Brennan, J., Takei, F., Wong, S., and Mager, D.L. (1995). Carbohydrate recognition by a natural killer cell receptor, Ly-49C. J. Biol. Chem. *270*, 9691–9694.

Brennan, J., Lemieux, S., Freeman, J., Mager, D., and Takei, F. (1996a). Heterogeneity among Ly49C NK cells: characterization of highly related receptors with differing functions and expression patterns. J. Exp. Med. *184*, 2085–2090.

Brennan, J., Mahon, G., Mager, D.L., Jefferies, W.A., and Takei, F. (1996b). Recognition of class I major histocompatibility complex molecules by Ly-49—Specificities and domain interactions. J. Exp. Med. *183*, 1553–1559.

Busch, D.H., Pilip, I.M., Vijh, S., and Pamer, E.G. (1998). Coordinate regulation of complex T cell populations responding to bacterial infection. Immunity *8*, 353–362.

Corral, L., Takizawa, H., Hanke, T., Jamieson, A.M., and Raulet, D.H. (1999). A new monoclonal antibody reactive with several Ly49 NK cell receptors mediates redirected lysis of target cells. Hybridoma, in press.

Correa, I., and Raulet, D.H. (1995). Binding of diverse peptides to MHC class I molecules inhibits target cell lysis by activated natural killer cells. Immunity *2*, 61–71.

Daniels, B.F., Karlhofer, F.M., Seaman, W.E., and Yokoyama, W.M. (1994a). A natural killer cell receptor specific for a major histocompatibility complex class I molecule. J. Exp. Med. *180*, 687–692.

Daniels, B.F., Nakamura, M.C., Rosen, S.D., Yokoyama, W.M., and Seaman, W.E. (1994b). Ly-49A, a receptor for H-2D^d, has a functional carbohydrate recognition domain. Immunity *1*, 785–792.

George, T., Yu, Y.Y.L., Liu, J., Davenport, C., Lemieux, S., Stoneman, E., Mathew, P.A., Kumar, V., and Bennett, M. (1997). Allorecognition by murine natural killer cells: lysis of T-lymphoblasts and rejection of bone marrow grafts. Immunol. Rev. *155*, 29–40.

George, T.C., Mason, L.H., Ortaldo, J.R., Kumar, V., and Bennett, M. (1999). Positive recognition of MHC class I molecules by the Ly49D receptor of murine NK cells. J. Immunol. *162*, 2035–2043.

Held, W., and Raulet, D.H. (1997). Ly49A transgenic mice provide evidence for a major histocompatibility complex-dependent education process in NK cell development. J. Exp. Med. *185*, 2079–2088.

Held, W., Roland, J., and Raulet, D.H. (1995). Allelic exclusion of Ly49 family genes encoding class I-MHC-specific receptors on NK cells. Nature *376*, 355–358.

Held, W., Cado, D., and Raulet, D.H. (1996a). Transgenic expression of the Ly49A natural killer cell receptor confers class I major histocompatibility complex (MHC)-specific inhibition and prevents bone marrow allograft rejection. J. Exp. Med. *184*, 2037–2041.

Held, W., Dorfman, J.R., Wu, M.-F., and Raulet, D.H. (1996b). Major histocompatibility complex class I-dependent skewing of the natural killer cell Ly49 receptor repertoire. Eur. J. Immunol. *26*, 2286–2292.

Hoglund, P., Sundbäck, J., Ollson-Alheim, M.Y., Johansson, M., Salcedo, M., Öhlén, C., Ljunggren, H.G., Sentman, C., and Karre, K. (1997). Host MHC class I gene control of NK cell specificity in the mouse. Immunol. Rev. *155*, 11–28.

Johansson, M.H., Hoglund, E., Nakamura, M.C., Ryan, J.C., and Hoglund, P. (1998). α1/α2 domains of H-2D^d, but not L^d, induce "missing self" reactivity in vivo—no effect of H-2L^d on protection against NK cells expressing the inhibitory receptor Ly49G2. Eur. J. Immunol. *28*, 4198–4206.

Kane, K. (1994). Ly-49 mediates EL4 lymphoma adhesion to isolated class I major histocompatibility complex molecules. J. Exp. Med. *179*, 1011–1015.

Karlhofer, F.M., Ribaudo, R.K., and Yokoyama, W.M. (1992). MHC class I alloantigen specificity of Ly-49$^+$ IL-2 activated natural killer cells. Nature *358*, 66–70.

Karlhofer, F.M., Hunziker, R., Reichlin, A., Margulies, D.H., and Yokoyama, W.M. (1994). Host MHC class I molecules modulate in vivo expression of a NK cell receptor. J. Immunol. *153*, 2407–2416.

Lian, R.H., Freeman, J.D., Mager, D.L., and Takei, F. (1998). Role of conserved glycosylation site unique to murine class I MHC in recognition by Ly-49 NK cell receptor. J. Immunol. *161*, 2301–2306.

Ljunggren, H.G., and Karre, K. (1990). In search of the 'missing self': MHC molecules and NK cell recognition. Immunol. Today *11*, 237–244.

Mason, L.H., Ortaldo, J.R., Young, H.A., Kumar, V., Bennett, M., and Anderson, S.K. (1995). Cloning and functional characteristics of murine LGL-1: a member of the Ly-49 gene family (Ly-49G2). J. Exp. Med. *182*, 293–303.

Mason, L.H., Anderson, S.K., Yokoyama, W.M., Smith, H.R.C., Winkler-Pickett, R., and Ortaldo, J.R. (1996). The Ly49D receptor activates murine natural killer cells. J. Exp. Med. *184*, 2119–2128.

Matsumoto, N., Ribaudo, R.K., Abastado, J.P., Margulies, D.H., and Yokoyama, W.M. (1998). The lectin like NK cell receptor Ly49A recognizes a carbohydrate independent epitope on its MHC ligand. Immunity *8*, 245–254.

McQueen, K.L., Freeman, J.D., Takei, F., and Mager, D. (1998). Localization of five new Ly49 genes, including three closely related to Ly49C. Immunogenetics *48*, 174–183.

Murphy, W.J., Kumar, V., Cope, J.C., and Bennett, M. (1990). An absence of T cells in murine bone marrow allografts leads to an increased susceptibility to rejection by natural killer cells and T cells. J. Immunol. *144*, 3305–3311.

Nagasawa, R., Gross, J., Kanagawa, O., Townsend, K., Lanier, L.L., Chiller, J., and Allison, J.P. (1987). Identification of a novel T cell surface disulfide-bonded dimer distinct from the α/β antigen receptor. J. Immunol. *138*, 815–824.

Nakamura, M.C., Linnemeyer, P.A., Niemi, E.C., Mason, L.H., Ortaldo, J.R., Ryan, J.C., and Seaman, W.E. (1999). Mouse Ly-49D recognizes H-2D(d) and activates natural killer cell cytotoxicity. J. Exp. Med. *189*, 493–500.

Orihuela, M., Margulies, D.H., and Yokoyama, W.M. (1996). The natural killer cell receptor Ly-49A recognizes a peptide-induced conformational determinant on its major histocompatibility complex class I ligand. Proc. Natl. Acad. Sci. USA *93*, 11792–11797.

Ortaldo, J.R., Winkler-Pickett, R., Mason, A.T., and Mason, L.H. (1998). The Ly-49 family: regulation of cytotoxicity and cytokine production in murine CD3$^+$ cells. J. Immunol. *160*, 1158–1165.

Perarnau, B., Saron, M.-F., Martin, B.R.S.M., Bervas, N., Ong, H., Soloski, M., Smith, A.G., Ure, J.M., Gairin, J.E., and Lemonnier, F.A. (1999). Single *H2K^b*, *H2D^b* and double *H2K^{b}D^b* knockout mice: peripheral CD8$^+$ T cell repertoire and anti-lymphocytic choriomeningitis virus cytolytic responses. Eur. J. Immunol. *29*, 1243–1252.

Pircher, H., Bürki, K., Lang, R., Hengartner, H., and Zinkernagel, R.M. (1989). Tolerance induction in double specific T-cell receptor transgenic mice varies with antigen. Nature *342*, 559–561.

Raulet, D.H., Held, W., Correa, I., Dorfman, J., Wu, M.-F., and Corral, L. (1997). Specificity, tolerance and developmental regulation of natural killer cells defined by expression of class I-specific Ly49 receptors. Immunol. Rev. *155*, 41–52.

Raziuddin, A., Longo, D., Mason, L., Ortaldo, J., and Murphy, W. (1996). Ly-49 G2$^+$ NK cells are responsible for mediating the rejection of H-2^b bone marrow allografts in mice. Intl. Immunol. *8*, 1833–1839.

Smith, H.R.C., Karlhofer, F.M., and Yokoyama, W.M. (1994). Ly-49 multigene family expressed by IL-2-activated NK cells. J. Immunol. *153*, 1068–1079.

Stoneman, E.R., Bennett, M., An, J., Chesnut, K.A., Wakeland, E.K., Scheerer, J.B., Siciliano, M.J., Kumar, V., and Mathew, P.A. (1995). Cloning and characterization of 5E6 (Ly49C), a receptor molecule expressed on a subset of murine natural killer cells. J. Exp. Med. *182*, 305–313.

Summers, M., and Smith, G. (1987). A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures (College Station, TX: Texas Agricultural Experiment Station).

Sundbäck, J., Nakamura, M.C., Waldenström, M., Niemi, E.C., Seaman, W.E., Ryan, J.C., and Kärre, K. (1998). The alpha2 domain of H-2D^{d} restricts the allelic specificity of the murine NK cell inhibitory receptor Ly-49A. J. Immunol. *160*, 5971–5978.

Takei, F., Brennan, J., and Mager, D.L. (1997). The Ly49 family: genes proteins and recognition of class I MHC. Immunol. Rev. *155*, 67–77.

Vance, R.E., and Raulet, D.H. (1998). Toward a quantitative analysis of the repertoire of class I MHC-specific inhibitory receptors on natural killer cells. Curr. Topics Microbiol. Immunol. *230*, 135–160.

Vance, R.E., Kraft, J.R., Altman, J.D., Jensen, P.E., and Raulet, D.H. (1998). Mouse CD94/NKG2A is a natural killer cell receptor for the nonclassical MHC class I molecule Qa-1^{b}. J. Exp. Med. *188*, 1841–1848.

Wong, S., Freeman, J.D., Kelleher, C., Mager, D., and Takei, F. (1991). Ly-49 multigene family. New members of a superfamily of type II membrane proteins with lectin-like domains. J. Immunol. *147*, 1417–1423.

Yokoyama, W.M. (1995). Natural killer cell receptors specific for major histocompatibility complex class I molecules. Proc. Natl. Acad. Sci. USA *92*, 3081–3085.

Yu, Y.Y., George, T., Dorfman, J., Roland, J., Kumar, V., and Bennett, M. (1996). The role of Ly49A and 5E6 (Ly49C) molecules in hybrid resistance mediated by murine natural killer cells against normal T cell blasts. Immunity *4*, 67–76.

Zijlstra, M., Bix, M., Simister, N.E., Loring, J.M., Raulet, D.H., and Jaenisch, R. (1990). β2-microglobulin deficient mice lack CD4 8^{+} cytolytic T cells. Nature *344*, 742–746.