Hyaluronan Recognition Mode of CD44 Revealed by Cross-saturation and Chemical Shift Perturbation Experiments*

Mitsuhiko Takeda‡‡, Hiroaki Terasawa‡, Masayoshi Sakakura‡, Yoshihiko Yamaguchi‡, Masahiro Kajiwara‡, Hirotoshi Kawashima‡, Masayuki Miyasaka‡, and Ichio Shimada‡‡‡

From the ‡Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, the ‡Japan Biological Information Research Center (JBIRC), Japan Biological Informatics Consortium (JBIC), Hatchobori, Chuo-ku, Tokyo 104-0032, Japan, the *Department of Medicinal Chemistry, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose-shi, Tokyo 204-8588, Japan, the ‡ Laboratory of Molecular and Cellular Recognition, Osaka University Graduate School of Medicine, 2-2, Yamada-oka, Suita 565-0871, Japan, and the **Biological Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), Aomi, Koto-ku, Tokyo 135-0064, Japan

CD44 is the main cell surface receptor for hyaluronic acid (HA) and contains a functional HA-binding domain (HABD) composed of a Link module with N- and C-terminal extensions. The contact residues of human CD44 HABD for HA have been determined by cross-saturation experiments and mapped on the topology of CD44 HABD, which we elucidated by NMR. The contact residues are distributed in both the consensus fold for the Link module superfamily and the additional structural elements consisting of the flanking regions. Interestingly, the contact residues exhibit small changes in chemical shift upon HA binding. In contrast, the residues with large chemical shift changes are localized in the C-terminal extension and the first α-helix and are generally inconsistent with the contact residues. These results suggest that, upon ligand binding, the C-terminal extension and the first α-helix undergo significant conformational changes, which may account for the broad ligand specificity of CD44 HABD.

CD44 is a type I transmembrane glycoprotein with diverse functions and is expressed on the surface of many cell types (1, 2). CD44 reportedly recognizes a variety of ligands (1, 2). The most investigated aspect of CD44 function is its ability to bind hyaluronic acid (HA), a major component of the extracellular matrix (1–3). HA is a very high molecular mass glycosaminoglycan, composed of a repeating disaccharide, D-glucuronic acid–N-acetyl-D-glucosamine (β1→3)-N-acetyl-D-glucosamine (β1→4) (4). The binding of CD44 to HA has been implicated in both cell adhesion to the extracellular matrix components and cellular signaling cascades (1, 2, 5). While HA exists as a high molecular mass polymer, HA fragments of various molecular sizes can be generated in vivo by a variety of mechanisms, and they exhibit different biological activities (6, 7). In addition to HA, CD44 interacts with the chondroitin sulfate (ChS) proteoglycans, serglycin (8–10), versican (11), and aggrecan (12). The CD44-binding elements on these proteoglycans are ChS side chains. These ChS proteoglycans may be involved in the adherence and activation of CD44-expressing cells.

CD44 has an N-terminal functional domain that interacts with the HA and ChS chains (11–13). This ligand recognition domain contains a homology region, termed the Link module (14, 15). Link modules are found in extracellular matrix molecules (link protein, aggrecan, versican, neurocan, and brevican) and the protein product of tumor necrosis factor-stimulated gene-6 (TSG-6). The HA-binding domains from Link module-containing proteins are divided into three subgroups as described in Ref. 16. Briefly, a single Link module, a Link module with N- and C-terminal extensions, and a pair of Link modules, which are sufficient for a high affinity interaction with HA, are classified as Types A, B, and C, respectively. In addition, a single Link module from TSG-6, belonging to Type A, interacts specifically with chondroitin 4-sulfate (Ch4S) but not with chondroitin 6-sulfate (Ch6S) (17). Bovine link protein, which contains the Type C Link module, binds to neither Ch4S nor Ch6S (18). On the other hand, CD44, which belongs to Type B, interacts with Ch4S, Ch6S, chondroitin, and other ChS chains (11, 19), suggesting that the HA-binding domain of CD44 has a broader ChS binding specificity than those of the TSG-6 and link proteins. The three-dimensional structure and ligand-binding site of Type A, i.e. the TSG-6 Link module, have been determined by NMR (20, 21). In contrast, no structural information about Type B, including CD44, is available.

In the present study, a combination of cross-saturation and chemical shift perturbation experiments has been applied to the CD44 HA-binding domain-HA complex to separate the direct intermolecular contacts from the effects due to ligand-induced conformational changes (22, 23). Significant conformational changes were observed in the C-terminal extension and the first α-helix, suggesting a correlation between the conformational changes and the broad ligand specificity of the CD44 HA-binding domain.

EXPERIMENTAL PROCEDURES

CD44 HABD and HA Sample Preparation—The human CD44 HABD gene (residues 21–178) was amplified by PCR from a cDNA encoding human CD44-IgG fusion protein (3, 24) and was cloned into the pET11a vector (Novagen) (24). CD44 HABD was expressed in Escherichia coli BL21 (DE3) codon Plus RP (Stratagene) without affinity tags and then

* This work was supported by a grant from the Japan New Energy and Industrial Technology Development Organization (NEDO) and Ministry of Economy, Trade and Industry (METI). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Tel./Fax: 81-3-3815-6549; E-mail: shimada@iw-nmr.e-u-tokyo.ac.jp.

‡‡ To whom correspondence should be addressed: Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Tel./Fax: 81-3-3815-6549; E-mail: shimada@iw-nmr.e-u-tokyo.ac.jp.

‡‡‡ To whom correspondence should be addressed: Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Tel./Fax: 81-3-3815-6549; E-mail: shimada@iw-nmr.e-u-tokyo.ac.jp.

‡‡‡‡‡ To whom correspondence should be addressed: Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Tel./Fax: 81-3-3815-6549; E-mail: shimada@iw-nmr.e-u-tokyo.ac.jp.

The abbreviations used are: HA, hyaluronic acid; ChS, chondroitin sulfate; TSG-6, tumor necrosis factor-stimulated gene-6; Ch4S, chondroitin 4-sulfate; Ch6S, chondroitin 6-sulfate; CS, cross-saturation; AMS, 4-acetoamide-4’-maleimidylisothiobene-2’-disulfonate 2Na; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single quantum coherence.

This paper is available on line at http://www.jbc.org
was purified as described previously (24). For the unlabeled protein, the transformed cells were grown in Luria-Bertani broth. Uniformly $^{15}$N- and $^{13}$C-labeled proteins were prepared by growing cells in M9 medium containing $^{15}$NH$_4$Cl (1 g/liter) and unlabeled or [U-$^{13}$C]glucose (2 g/liter) supplemented with Celtone-DN powder (1 g/liter; Martek). Similarly, the uniformly $^{15}$N-$^{13}$C-labeled protein was prepared using M9 medium in 99.9% H$_2$O containing $^{15}$NH$_4$Cl (1 g/liter) and [U-$^{13}$C]glucose (98%; 2 g/liter) supplemented with Celtone-DN powder (1 g/liter). In addition, several selective $^{15}$N labelings of one type of amino acid were carried out to provide starting points for the sequential assignments (25).

The purified and lyophilized protein was dissolved in 6 M guanidine, 100 mM Tris, pH 8.0, and 150 mM NaCl and was dialyzed against H$_2$O/2H$_2$O (9/1) buffer containing 50 mM potassium phosphate, pH 6.7, and 150 mM NaCl. In cross-saturation (CS) experiments, the proton content of the solvent was lowered to 20% to suppress the spin diffusion caused by the exchangeable amide protons in CD44 HABD. Sample concentrations were 0.3 mM (uniformly $^{15}$N-labeled for titration experiments), 2.0 mM (13C/15N-labeled for sequential assignments), and 12 mM (H$_2$O/2H$_2$O-labeled for CS experiments).

HA$_{34}$ and HA$_{34}$ were prepared as described previously (26). A stock solution of 60 mM HA$_{34}$ or 14 mM HA$_{34}$ was added to the CD44 HABD samples.

Characterization of CD44 HABD—The redox states of the refolded CD44 HABD were analyzed by the 4-acetoamide-4-maleimidylstilbene-2,2'-disulfonate 2Na (AMS) modification and SDS-PAGE, as described by Uchida et al. (27). Surface plasmon resonance measurements were performed using a BIAcore 2000 instrument (Biacore AB). Biotinylated HA was prepared as described previously (28) and was immobilized onto the sensor chip SA (Biacore AB), resulting in a signal of 100 resonance units. The binding assay was performed in running buffer (10 mM sodium phosphate, pH 7.0, and 150 mM NaCl) at flow rate of 10 $\mu$l/min, using serial dilutions of CD44 HABD in the 5–400 nM range. The dissociation constants ($K_d$) were obtained from the steady-state curve fitting analysis, using the BIAevaluation software, version 3.0 (Biacore AB).

NMR Spectroscopy—All NMR experiments were performed at 25°C on Bruker DRX 400, Avance 500, or Avance 600 spectrometers. The sequential assignments of the $^1$H, $^{15}$N, $^{13}$C, and $^{13}$C labels of CD44 HABD in the unbound and HA$_{34}$ bound states were achieved by the sets of three-dimensional triple resonance experiments (HNCA, HN(CO)CA, CBCA(CO)NH, and HN(CO)CA) (29). All spectra were processed by NMRPipe (30), and data analysis was assisted by ANSIG (31). The NOE data analysis was assisted by the Sparky software (32). In the HA$_{34}$-bound state, the NOE cross-peak assignments were limited by signal broadening, due to the high molecular mass of the complex.

In the measurement of the diffusion coefficients of CD44 HABD by pulsed-field gradient, the pulsed-field gradient longitudinal eddy-current delay pulse sequence was used (33). The magnitude of the signal from the methyl protons at $\delta$ 0.8 ppm was used for determining the signal intensities of the protein.

In the CS experiments, the molar ratio of HA$_{34}$ to CD44 HABD was set to 1.7:1. The CS experiments were performed using the pulse scheme (23), with a minor modification from the transverse relaxation-optimized polarization-transfer-type coherence transfer to the insensitive nuclei enhanced by polarization transfer-type one (34). The saturation for the HA protons was made using continuous-wave RF irradiation. The saturation frequency was set at 1.9 ppm (methyl protons) or 3.5 ppm (ring protons). The reduction ratios of the signal intensities were calculated with the average of the results from these two experiments. The results of significantly overlapped residue were excluded. Each measurement time was 48 h, with a relaxer 60% supplied and a saturation time of 2.0 s.

RESULTS

Characterization of CD44 HABD—The boundaries of CD44 HABD were defined by analyses using truncations and site-directed mutagenesis (13). We refer to the Link module of CD44 (residues 32–124) as CD44 LM, and the minimum region required for HA binding (residues 21–178) as CD44 HABD (HA-binding domain), as shown in Fig. 1A. Recombinant CD44 HABD was prepared based on the method developed by Banerji et al. (24) with minor modifications (see “Experimental Procedures”). As shown in Fig. 1B, the $^1$H-$^{15}$N HSQC spectrum of CD44 HABD is well dispersed, indicating that CD44 HABD is a structural domain. Banerji et al. (24) reported that CD44 HABD retains its reactivity with several conformation-sensitive antibodies and has a similar HA binding activity to that of CD44 molecules expressed in eukaryotic cells. To provide further evidence for the proper folding and function of CD44 HABD, we performed additional analyses. First, to examine the redox states of the six cysteine residues with CD44 HABD, we used AMS, which binds to the sulphydryl (SH) group (27). CD44 HABD showed an upward shift in its position on an SDS gel when it was treated with AMS in the presence of dithiothreitol. No such shift was observed in the absence of dithiothreitol (data not shown). These results indicated that CD44 HABD contains three disulfide bonds and no free SH group. Second, to evaluate the HA binding activity of CD44 HABD, surface plasmon resonance measurements were employed. The results showed that CD44 HABD has HA binding activity, and the dissociation constant ($K_d$) for CD44 HABD toward the HA 250-mer is 2.7 × 10$^{-5}$ M (the $K_d$ for the monomeric CD44Rg protein, expressed in ldl-D cells and deglycosylated, toward the HA 60-mer is 5 × 10$^{-6}$ M (35)). Finally, the in vitro refolding of CD44 HABD using DeBC, a disulfide isomerase with chaperone activity (36), was achieved. By SDS-PAGE, the resultant product was shown to be identical to the protein prepared by the reported method, suggesting that CD44 HABD adopts a native fold. The details of the refolding method using DeBC will be reported elsewhere. Taken together, these results strongly support the idea that CD44 HABD retains the native fold and the proper functions.

Topology of CD44 HABD—As a first step toward elucidating the molecular basis for the ligand recognition by CD44 HABD, we investigated the topology of CD44 HABD in an unbound state. By three-dimensional NMR analyses, we assigned the main chain resonance of CD44 HABD (29) (see “Experimental Procedures”). On the basis of both the NOE connectivities observed among the main chain C$^\alpha$ and H$^N$ resonances and the chemical shift index analysis (37), we determined the secondary structure and the topology of CD44 HABD (Fig. 1, D and E). CD44 HABD was identified as a structural domain, comprising two $\alpha$-helices (a1, 46–56; a2, 63–71) and two $\beta$-sheets. Sheet I is composed of five strands: $\beta_a$ (residues 22–27), $\beta_b$ (34–39), $\beta_c$ (116–118), $\beta_d$ (144–148), and $\beta_f$ (155–158). Sheet II is a two-stranded antiparallel $\beta$-sheet: $\beta_e$ (79–82) and $\beta_d$ (85–88). A comparison of the topology with that of TSG-6 indicated that CD44 LM adopts the consensus fold for the Link module superfamily (20). It should be noted that the N- and C-terminal extensions form three additional $\beta$-strands ($\beta_b$, $\beta_e$, and $\beta_f$) to sheet I of the consensus fold (Fig. 1E, green).

Cross-saturation Experiments—To identify the residues of CD44 HABD that contact HA, CS experiments were performed (23) (Fig. 2A). HA oligosaccharides of defined lengths were prepared, as reported by Tawada et al. (26). HA with an average molecular mass of 6.9 kDa (termed HA$_{34}$) was selected as a saturation-donating partner from several HA oligomers for the following reasons. First, HA$_{34}$ is involved in many biological processes, such as the induction of Fas expression and the Fas-mediated apoptosis of synovial cells (38), the activation of integrins on colon cancer cells (39), and the enhancements of CD44 cleavage from tumor cells and their motility (7). Second, HA$_{34}$ is likely to fully cover the HA-binding surface of CD44 HABD. Finally, the saturation can be transferred more effectively from HA$_{34}$ to the labeled CD44 HABD than from short-chain HA oligosaccharides (e.g. the HA hexasaccharide, termed...
FIG. 1. Sequence, NMR spectra, and topology of CD44 HABD. A, schematic diagram showing the sequence of the expressed CD44 HABD. The residual N-terminal sequence derived from the multiple cloning site is designated by lowercase letters. B and C, two-dimensional $^1$H-$^1$H HSQC spectra of uniformly $^{15}$N-labeled CD44 HABD in the unbound (B) and HA$_{34}$/H11011$_{34}$ bound (C) states. The cross-peaks are labeled with the one-letter codes for amino acids and the residue numbers. D, interstrand backbone NOE connectivities of CD44 HABD. Interstrand NOEs observed in three-dimensional $^{15}$N-edited and two-dimensional NOESY spectra are marked by arrows. E, topology of CD44 HABD. Strands and helices are represented by arrows and cylinders, respectively. The Link module region is depicted with a black line, and the N- and C-terminal extensions are shown with green lines.
HA_{2}) in the complex. Pulsed-field gradient NMR experiments (33) generated diffusion coefficients of CD44 HABD in complex with HA_{24} and HA_{6} of 0.6 \times 10^{-10} \text{ m}^2/\text{s} and 1.1 \times 10^{-10} \text{ m}^2/\text{s} at 25 °C, respectively, indicating that the rotational correlation time of the HA_{24} bound complex is longer than that of the HA_{6} bound complex. Indeed, the CS phenomena were efficiently observed for the CD44 HABD-HA_{24} complex.

**HA-binding Site**—The main chain resonances of CD44 HABD in complex with HA_{24} were assigned by three-dimensional NMR analyses (Fig. 1C). The rf irradiation applied to the complex resulted in selective intensity losses for the CD44 HABD resonances by the CS phenomena (Fig. 2B). On the basis of the spectra for the complex with and without irradiation, the reduction ratios of the peak intensities were calculated (Fig. 3A). The residues strongly affected by the irradiation were mapped on the topology of CD44 HABD (Fig. 3D). The affected residues are distributed on both the consensus fold and the additional structural elements consisting of the flanking regions.

**Conformational Changes upon HA Binding**—The weighted averaged {^1}H and {^{15}}N chemical shift changes between the HA_{24} bound and unbound states of CD44 HABD are plotted in Fig. 3B. The residues with large chemical shift differences are also mapped on the topology of CD44 HABD. Interestingly, these residues are mostly localized on the C-terminal extension and the a1 helix, designated as the Shift region (Fig. 3D, colored in *light yellow*), and they are generally inconsistent with the contact residues. Significant chemical shift changes upon HA binding are observed in the {^{15}}Co resonance of the residues in the Shift region (data not shown), supporting the idea that the conformational changes upon HA binding occur in this region.

The amino acids of CD44 HABD that mediate the interactions with HA have been investigated by site directed mutagenesis (13, 40). Forty residues have been shown to be important in HA binding (Fig. 3D). Twelve residues (*i.e.* Arg^{29}, Lys^{38}, Arg^{41}, Tyr^{42}, Lys^{68}, Arg^{78}, Tyr^{79}, Asn^{100}, Asn^{101}, Tyr^{105}, Arg^{150}, and Arg^{154}) are located either within or in close proximity to the HA-binding site. On the other hand, Lys^{158} and Arg^{162} are apart from the binding site. We cannot completely rule out the possibility that CS effect is attenuated by the flexibility in the Shift region. Therefore, CS experiments using side chain resonances are currently under way to examine the importance of these residues in HA binding.

**DISCUSSION**

The HA-binding site of CD44 HABD was successfully identified by the cross-saturation experiments in the present study. The N-linked glycosylation sites of CD44 HABD, previously shown to be involved in the regulation of HA binding (41), are indicated in Fig. 3D. Two glycosylation sites, Asn^{67} and Asn^{100}, are located within the HA-binding site.

Kahmann et al. (21) reported that Lys^{11}, Tyr^{49}, Asn^{67}, Phe^{70}, Lys^{72}, and Tyr^{78} of the TSG-6 Link module could be involved directly in HA binding. The HA contact residues of CD44 and the putative binding residues of TSG-6 are shown in the alignment of these sequences, based on the secondary structure elements (Fig. 3C). Although the HA-binding sites of CD44 and TSG-6 are likely to overlap in the β4-β5 region (in TSG-6), CD44 may have a larger HA-binding site than that of TSG-6.

It was suggested that the HABD of CD44 has broader ligand specificity than those of the TSG-6 and link proteins (11, 17–19). We propose that the broad specificity of CD44 HABD can be accounted for by the ligand-induced conformational changes in the Shift region. It is likely that the HA-binding site of CD44 HABD can be divided into two regions: one formed by the Link module region, which may play a role in comparatively specific ligand-recognition, and the other is the Shift region, which may adjust to accommodate structurally different ligands. In our preliminary studies, some of the Shift region signals showed different chemical shift changes upon chondroitin binding, compared with HA binding, in the {^1}H-{^{15}}N HSQC spectra (data not shown), suggesting that the Shift region adopts a somewhat different conformation in the chondroitin-bound complex from that of the HA-bound complex.

Since protein-carbohydrate interactions are implicated in many biological events, from development to tumorigenesis, understanding carbohydrate recognition at an atomic resolution is of utmost importance (42). Chemical shift perturbation is a widely used NMR method to map protein interfaces. However, in this report, chemical shift perturbation would not be an appropriate method to identify the ligand-binding site of CD44 HABD because of the small chemical shift changes at the contact site and the presence of significant conformational changes in the flanking regions. Several chemical shift perturbation experiments for protein-carbohydrate complexes have been reported (21, 43–46). A common feature of these experiments is that the chemical shift perturbations of the protons on the putative ligand-binding sites are at most 0.35 ppm. This feature may reflect the fact that the common molecular scaffold of carbohydrates has no aromatic group that would cause magnetic anisotropies (Fig. 3E). On the other hand, the efficiency of CS does not depend on the presence of an aromatic group in the ligand molecules. Therefore, the CS method is especially helpful to determine interfaces of proteins with carbohydrates that have no source of ring current shifts.

In this study, we have identified the topology and the ligand recognition mode of CD44 HABD. The topology will be helpful for modeling other HA-binding domains belonging to the Type B Link module family. A combination of the CS and chemical shift perturbation methods may provide a useful means to separate the direct intermolecular contacts from the effects due to ligand-induced conformational changes.

**Acknowledgments**—We are grateful to Dr. B. Seed for providing the cDNA encoding human CD44-IgG fusion protein and Drs. Akira Asari,
FIG. 3. Comparison between cross-saturation and chemical shift perturbation experiments. A, plots of the reduction ratios of the signal intensities originating from the backbone amide groups, with and without presaturation, along with the determined secondary structure. The residues with reduction ratios >0.2 are colored red, and those with reduction ratios within the 0.1–0.2 range are orange. B, plots of the weighted averaged 1H and 15N chemical shift changes calculated with the function \( \Delta = \left( \delta_{\text{H}} + 0.25 \delta_{\text{N}} \right) / 2 \). Residues with chemical shift changes \( \Delta > 0.6 \) ppm are colored cyan. C, an alignment of the CD44 HABD and TSG-6 Link module sequences, based on the secondary structure elements. The CD44 HABD residues affected by CS experiments are colored as described in the legend to A. Amino acids that could be involved directly in HA binding by the TSG-6 Link module (as described by Kahmann et al. (21)) are colored purple. Amino acid insertions are indicated with triangles. D, mapping of the residues affected in the CS and chemical shift perturbation experiments on the topology of CD44 HABD. The color usage is as described above. Ile26, Ala98, and Asn149 were affected in both experiments and are doubly colored. The Shift region, which is covered with residues that are strongly affected in the chemical shift perturbation experiments, is shown in light yellow. The residues of the N-glycosylation sites are labeled with brown circles (41). The amino acids that have been investigated by site-directed mutagenesis are labeled with stars (13, 40). E, schematic representation of the chemical shift changes and CS effects due to sugar and aromatic rings.
REFERENCES