Molecular Cell, Vol. 13, 483–496, February 27, 2004, Copyright ©2004 by Cell Press

Structure of the Regulatory Hyaluronan Binding Domain in the Inflammatory Leukocyte Homing Receptor CD44

Peter Teriete,[1,3,6] Suneale Banerji,[4,6] Martin Noble,[2,3,6] Charles D. Blundell,[1,3] Alan J. Wright,[1,3] Andrew R. Pickford,[3] Edward Lowe,[2,3] David J. Mahoney,[1] Markku I. Tammi,[5] Jan D. Kahmann,[1,3] Iain D. Campbell,[3] Anthony J. Day,[1,*] and David G. Jackson[4,*]

[1]Medical Research Council Immunochemistry Unit and
[2]Laboratory of Molecular Biophysics
[3]Department of Biochemistry
University of Oxford
South Parks Road
Oxford OX1 3QU
[4]Medical Research Council Human Immunology Unit
Weatherall Institute of Molecular Medicine
John Radcliffe Hospital
Headington
Oxford OX3 9DS
United Kingdom
[5]Department of Anatomy
University of Kuopio
FIN-70211 Kuopio
Finland

Summary

Adhesive interactions involving CD44, the cell surface receptor for hyaluronan, underlie fundamental processes such as inflammatory leukocyte homing and tumor metastasis. Regulation of such events is critical and appears to be effected by changes in CD44 N-glycosylation that switch the receptor "on" or "off" under appropriate circumstances. How altered glycosylation influences binding of hyaluronan to the lectin-like Link module in CD44 is unclear, although evidence suggests additional flanking sequences peculiar to CD44 may be involved. Here we show using X-ray crystallography and NMR spectroscopy that these sequences form a lobular extension to the Link module, creating an enlarged HA binding domain and a formerly unidentified protein fold. Moreover, the disposition of key N-glycosylation sites reveals how specific sugar chains could alter both the affinity and avidity of CD44 HA binding. Our results provide the necessary structural framework for understanding the diverse functions of CD44 and developing novel therapeutic strategies.

Introduction

Cellular interactions with the extracellular matrix mucopolysaccharide hyaluronan (HA) are fundamentally important for cell migration during tissue morphogenesis, leukocyte trafficking, and tumor metastasis (Tammi et al., 2002; Toole et al., 2002). The major receptor mediating these processes, CD44, is an integral membrane glycoprotein belonging to the Link module superfamily, whose members include the matrix proteoglycans aggrecan, versican, and cartilage link protein (Day and Prestwich, 2002). Expressed on a variety of different cell types, CD44 plays a pivotal role on leukocytes as a receptor for transendothelial migration at sites of inflammation (Lesley et al., 1997). In a fashion reminiscent of the integrins, CD44 is functionally inactive on resting T cells and monocytes, and binds HA only after T cell receptor triggering or after activation by proinflammatory cytokines including TNF-α and INF-γ (reviewed in Pure and Cuff, 2001). In this way, discrete populations of cells in peripheral blood are directed to transmigrate at inflammatory sites, where they recognize and respond to HA displayed on the surface of activated endothelial cells (see Mohamadzadeh et al., 1998) as evidenced in mouse models of arthritis (Mikecz et al., 1995), autoimmune diabetes (Weiss et al., 2000), and asthma (Katoh et al., 2003). Other important functions of leukocyte CD44 relating to inflammation include the binding of small HA fragments and the transduction of signals that trigger activation of NFκB, and synthesis of chemokines (reviewed in Pure and Cuff, 2001; Noble, 2002).

Most HA binding proteins interact with ligand via a Link module (Day and Prestwich, 2002). The Link domain is approximately 100 residues in length, comprising two antiparallel β sheets and two α helices, stabilized by two highly conserved disulfide bridges, and is structurally related to the C-type lectins (Blundell et al., 2003; Kohda et al., 1996). In the Link module from TSG-6, which forms an independently folded structure sufficient for high-affinity binding (Kahmann et al., 2000; Kohda et al., 1996), the HA-interaction surface is comprised of basic and aromatic amino acids brought together from non-contiguous regions, generally on loops between elements of secondary structure (Mahoney et al., 2001b). However, there is clear evidence that the HA binding domain of CD44 is more complex, involving the additional contribution of sequences outside the consensus Link module just described. First, constructs comprising only the immediate Link module expressed in *E. coli* fail to refold and are functionally inactive (Banerji et al., 1998); the CD44 primary sequence encodes an extra pair of cysteine residues, located in the N- and C-terminal extensions, that is predicted to form a third disulfide linkage essential for stabilizing the HA binding domain. Second, site-directed mutagenesis indicates that basic amino acids in the flanking sequences participate in HA binding (Peach et al., 1993), in addition to those located in the Link module (Bajorath et al., 1998). Third, mapping of the epitopes for mAbs that block HA binding indicates that the cluster of basic residues located in the C-terminal extension is in close proximity to a region within the first 22 amino acids of the mature protein (Liao et al., 1995). These and other observations have led to the hypothesis that CD44 contains an extended HA binding domain that is specially adapted for the regulation of ligand binding (Jackson et al., 2001; Jackson, 2003).

One of the mechanisms by which CD44 is triggered

*Correspondence: ajday@bioch.ox.ac.uk (A.J.D); djackson@hammer.imm.ox.ac.uk (D.G.J.)

[6]These authors contributed equally to this work.

to bind HA following cell activation involves differential glycosylation of the receptor, specifically, the remodeling of N- and possibly O-linked carbohydrates (Levesque and Haynes, 1999). Although the precise structure of the CD44 glycan chains and the details of their remodeling are not completely clear, some important features are recognized. For example, studies with human myeloid cell lines and mouse lymphoid cell lines (Lesley et al., 1995) have shown that HA binding can be triggered by enzymatic removal of terminal sialic acid attached to N-linked glycans on CD44 by treatment with N-glycanase or by inhibition of biosynthesis with tunicamycin (Katoh et al., 1995). Moreover, the induction of HA binding by TNFα and phorbol ester, appears to be mediated by activation of an endogenous sialidase (Gee et al., 2003; Katoh et al., 1999; Kincade et al., 1997)—a likely physiological regulator of CD44 function. Five conserved N-glycosylation sites (Asn25, Asn57, Asn100, Asn110, and Asn120) are present within the N-terminal 150 residues of CD44, and the individual mutation of two of these, Asn25 and Asn120, has been shown to convert CD44 from an inducible to a constitutively active state in murine lymphoma cell lines (English et al., 1998). However, without knowledge of the 3D structure for CD44, it has not been possible to determine where these regulatory N-glycans lie in relation to the ligand binding site or precisely how they might influence the interaction with HA. In this manuscript, we have elucidated the 3D structure of the HA binding domain of CD44 by X-ray crystallography and NMR spectroscopy and have mapped the position of the HA-interaction surface. This has provided new insight into the mechanisms by which CD44 is regulated in vivo.

Results and Discussion

Expression and Characterization of the CD44 HA Binding Domain

For structural studies, we used soluble nonglycosylated CD44 HA binding domains (HABD) expressed in *E. coli* from the construct $CD44^{20\text{-}178}$ (identical to that described in Banerji et al. [1998] but without the N-terminal histidine tag) which encompasses the Link module together with its N- and C-terminal flanking sequences. We previously showed that the $CD44^{20\text{-}178}$ HABD is very similar to CD44Fc in terms of reactivity with the conformation-sensitive mAbs BRIC-235 and F10.44.2, indicating that the nonglycosylated protein expressed in *E. coli* shares a common fold with the fully glycosylated CD44 ectodomain expressed in human 293T fibroblasts (Banerji et al., 1999; data not shown).

Ligand binding assays revealed that $CD44^{20\text{-}178}$ binds biotinylated-HA (bHA) in a concentration-dependent manner that is similar to that of CD44Fc (Figure 1A). The observed differences in the binding curves for the two proteins may be attributed to avidity effects, since CD44Fc is a disulfide-linked dimer (Aruffo et al., 1990) whereas $CD44^{20\text{-}178}$ is monomeric. The plate binding assay also confirmed the HA binding specificity of $CD44^{20\text{-}178}$ (Aruffo et al., 1990) inasmuch as no competition was observed even with high concentrations of ultrapure C4S and C6S (Figure 1B).

Last, competitive binding analyses with HA oligosaccharides showed that the minimum effective unit is a 6-mer, but that 8- to 12-mers are more efficient competitors (Figure 1C). This confirms previous reports that the minimum size of oligosaccharide required to displace high M_r HA from cell surface CD44 is between 6 and 10 saccharide units (i.e., HA_6 or HA_{10}), depending on the cell background (see Lesley et al., 2000). Hence, the bacterially expressed, nonglycosylated $CD44^{20\text{-}178}$ molecule retains most or all of the functional characteristics of native glycosylated CD44.

Determination of the CD44 HABD Structure by NMR and X-Ray Crystallography

To investigate the structure of the CD44 HABD and its interaction with HA in solution, we applied the twin strategies of NMR spectroscopy and X-ray crystallography.

The NMR structure of $CD44^{20\text{-}178}$ was based on a total of 2168 nuclear Overhauser effect (NOE)-derived interproton distances (see Supplemental Figure S1 at http://www.molecule.org/cgi/content/full/13/4/483/DC1), 47 dihedral angle (ϕ) restraints, 56 H bonds between elements of secondary structure, and 157 $^{13}C^{\alpha}$ and 146 $^{13}C^{\beta}$ chemical shift values (see Table 1). The H bonds were determined from the observation of 37 slowly exchanging amide protons in regions of secondary structure and/or from NOE connectivities (see Supplemental Figure S2). In the final family of 20 structures (Figure 2A), residues 20–169 are generally well defined with an rmsd for backbone atoms of 1.21Å from the mean structure. Amino acids 170–178 (not shown) are significantly more disordered (see Table 1 and Figure 3A), and the high rmsd values observed in this region are consistent with the low number of NOEs identified per residue (Supplemental Figure S1). Furthermore, measurement of the $^{15}N\text{-}\{^{1}H\}$ NOE (Figure 3) revealed these amino acids to be dynamic in nature, consistent with their location in the flexible membrane-proximal stalk region of CD44 that is not directly involved in HA binding. NMR experiments conducted on $CD44^{20\text{-}178}$ samples with or without 150 mM NaCl gave identical spectra, indicating that the solution structure of the CD44 HABD determined here is likely to be representative of that found under physiological conditions.

The structure determined by X-ray crystallography was generated from both nonderivatized and selenomethione-derivatized (SeMet) $CD44^{20\text{-}178}$ (see Experimental Procedures). The 2-fold averaged SeMet single wavelength anomalous dispersion (SAD) map was readily interpretable for all but the last 9 amino acids (Figure 2D), and the resulting structure of residues Ala20-Tyr169 has been refined to an R factor of 18.4% at 2.2 Å resolution (see Table 2). Therefore, the crystal and solution structures of $CD44^{20\text{-}178}$ (determined independently) both show that residues 21–169 of human CD44 (i.e., the N-terminal 149 amino acids of the mature protein) constitute a structural domain.

The crystal and NMR structures of $CD44^{20\text{-}178}$ are very similar (Figure 2), with close agreement over regions of secondary structure. Superposition reveals an rmsd of 1.99 Å for 128 closely matching C^{α} atoms determined using O. Such a global comparison provides a rather poor index of their high degree of local similarity as can be seen from comparisons of interresidual contacts in

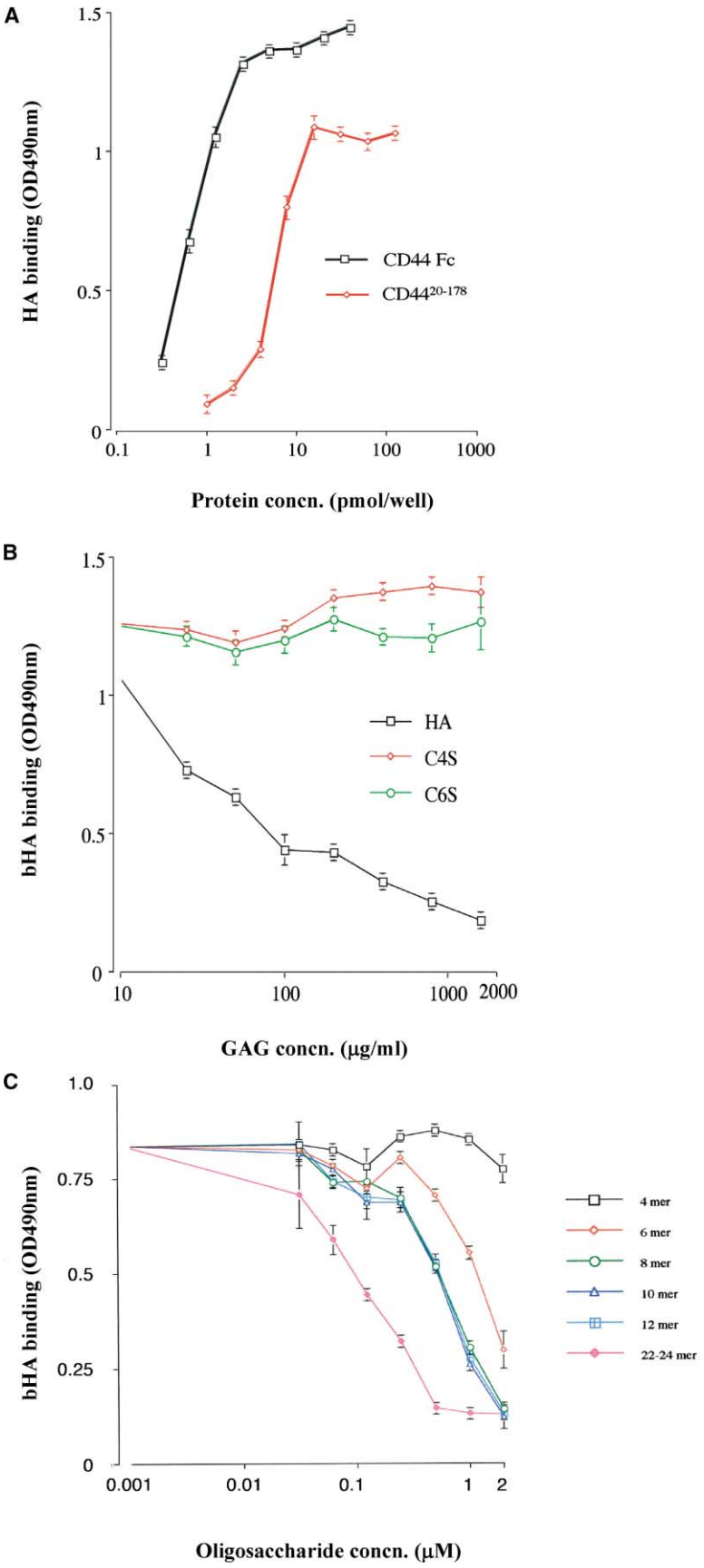

Figure 1. Hyaluronan Binding Characteristics of Nonglycosylated CD44$^{20-178}$

(A) Comparison of biotinylated-HA (bHA) binding to microtiter plates coated with either glycosylated (CD44Fc) or nonglycosylated CD44$^{20-178}$.

(B and C) Competition of bHA binding by unlabeled HA compared to the sulfated glycosaminoglycans C4S and C6S (B) or HA oligosaccharides of defined length (C). All values are plotted as mean ± SEM of triplicate determinations.

Table 1. NMR Restraints and Structural Statistics

NOE restraints	2168
Intraresidue	949
Sequential	476
Medium range (I < 5)	190
Long range (I ≥ 5)	553
Average number of NOE/residue	13.6
Dihedral angles (ϕ)	47
Hydrogen bonds	56
Total number of restraints	2271
Violations[a]	0
Distance restraints violated in single structure (>0.3 Å)	0
Distance restraints violated in all best 20 structures (>0.1 Å)	0
Dihedral angle restraints (>5°)	0
Precision[b]	
Backbone heavy atoms (Å)	0.612
All heavy atoms (Å)	1.129
Structure quality by PROCHECK[c]	
Residues in most favored regions	57.0%
Residues in additional allowed regions	34.4%
Residues in generously allowed regions	6.7%
Residues in disallowed regions	1.9%

[a] Values are derived from the 20 lowest energy structures.
[b] The average rmsd for the structure family was calculated by superimposing each of the 20 structures onto the coordinate set with the lowest target function. This superposition was over backbone Ca, C, N, and O atoms of residues 22–25, 28–30, 33–39, 46–56, 58–60, 62–72, 79–82, 85–88, 105–107, 115–120, 126–128, 143–149, 153–159.
[3] Laskowski et al., 1996.

the crystal structure, with interresidual NOEs in the NMR structure (see Supplemental Data). The near complete overlap of these values is testimony to the excellent agreement of the two techniques (see Supplemental Figure S3). However, given the greater precision of a high-resolution crystal structure compared with that obtained by NMR, especially with respect to details of side chain conformation, we have interpreted functional data below (including NMR shift mapping) on the basis of the X-ray structure.

Distinctive Features of the CD44 HABD

A search of the protein database using DALI (Holm and Sander, 1993) indicates that, while related to the Link module (and hence C-type lectins; Blundell et al., 2003), the structure of the CD44 HABD overall represents a previously unrecognized fold, i.e., a novel elaboration of the C-type lectin fold that is substantially larger than the HABD of TSG-6. Inspection of the crystal and NMR-derived structures for $CD44^{20\text{-}178}$ reveals the HABD to comprise two orthogonally disposed α helices and ten β strands organized into a single long, curving β sheet (Figures 2C and 2D; Supplemental Figure S2). In the recently reported secondary structure for the CD44 HABD (Takeda et al., 2003) three of these β stands are missing (e.g., β2, β5, and β7), and two unconnected β sheets (comprising two and five strands, respectively) were predicted. This incorrect secondary structure topology and lack of any 3D structural information made correct analysis of their ligand binding experiments impossible.

In the tertiary structures described here, the Link module of CD44 (residues 32–120) is very similar to that of TSG-6 (Figure 4), being comprised of two α helices and two triple-stranded antiparallel β sheets, SI (β strands 1, 2, and 6) and SII (β strands 3, 4, and 5), joined by H bonds between β3 and β6; disulfide bridges connect α1 to β6 and the loop following α2 to the long loop between β4 and β5 (Figure 2). However, in CD44, the regions flanking the Link module provide an additional four β strands (denoted 0, 7, 8, and 9) that extend the SI sheet. Specifically, the N-terminal sequence (residues 20–31) forms β0, a discontinuous strand interrupted by a β bulge, which connects to β1 of the Link module (see Supplemental Figure S2). The C-terminal extension forms three β strands (7, 8, and 9), where β7 (residues 126–128) and β8 (residues 144–148) partner residues 29–31 and 21–25 of β0, respectively, and β9 (153–159) is connected to β8 (143–149). Thus, the two extensions, which are linked by a disulfide bridge between Cys28 and Cys129 (see Figure 4E), together form an additional structural lobe in intimate contact with the Link module (Figure 4C). The large size of this lobe can be readily appreciated by comparisons of the surfaces presented by the CD44 and TSG-6 structures (Figures 4B and 4C).

Residues 21–169 that define the CD44 HABD are generally well ordered. The ^{15}N-{^{1}H} NOE experiment and crystallographic B factors (Figure 3) reveal that some of the loop regions are highly dynamic. One of these (residues 90–95) is found on the long, hook-shaped segment between β4 and β5. In the crystal structure this loop contains two very short β strands that generate a Greek-key architecture, with a half turn of helix between them (see Figures 2 and 4). The flexibility of this region, with no-NOE/NOE ratios close to zero and high temperature factors (Figures 3A and 3B), is probably due to *cis*/*trans* isomerization of prolines 89 and 93. Other highly dynamic residues are Asn110, a site for N-glycosylation in the flexible β5-β6 loop, and Glu126, the first residue of the β7 strand in the C-terminal extension (Figure 3B). Flexibility of the latter residue probably derives from *cis*/*trans* isomerism of neighboring prolines 124 and 125.

Alignment of the CD44 amino acid sequence with that of the lymphatic hyaluronan receptor LYVE-1 reveals that both proteins have additional conserved cysteines in the N- and C-terminal regions flanking their Link modules (Figure 4A). This suggests the LYVE-1 HABD may also have an extended Link module structure. However, the similarity with CD44 breaks down beyond strand β7 and LYVE-1 lacks the series of alternating hydrophobic and hydrophilic amino acids ($I^{144}TITI^{148}$) that form β8. Further structural studies will be required to resolve these issues.

Mapping the HA Binding Surface

Previously, a study using site-directed mutagenesis identified four key residues within the CD44 Link module (Arg41, Tyr42, Arg78, and Tyr79) as critical for the interaction with HA and a further five (Lys38, Lys68, Asn100, Asn101, and Tyr105) as functionally important (Bajorath et al., 1998). Mapping of these amino acids onto the CD44 HABD structure (Figure 5) reveals that all except Lys68 (which is on the opposite face of the protein) come together as a contiguous linear patch in a similar

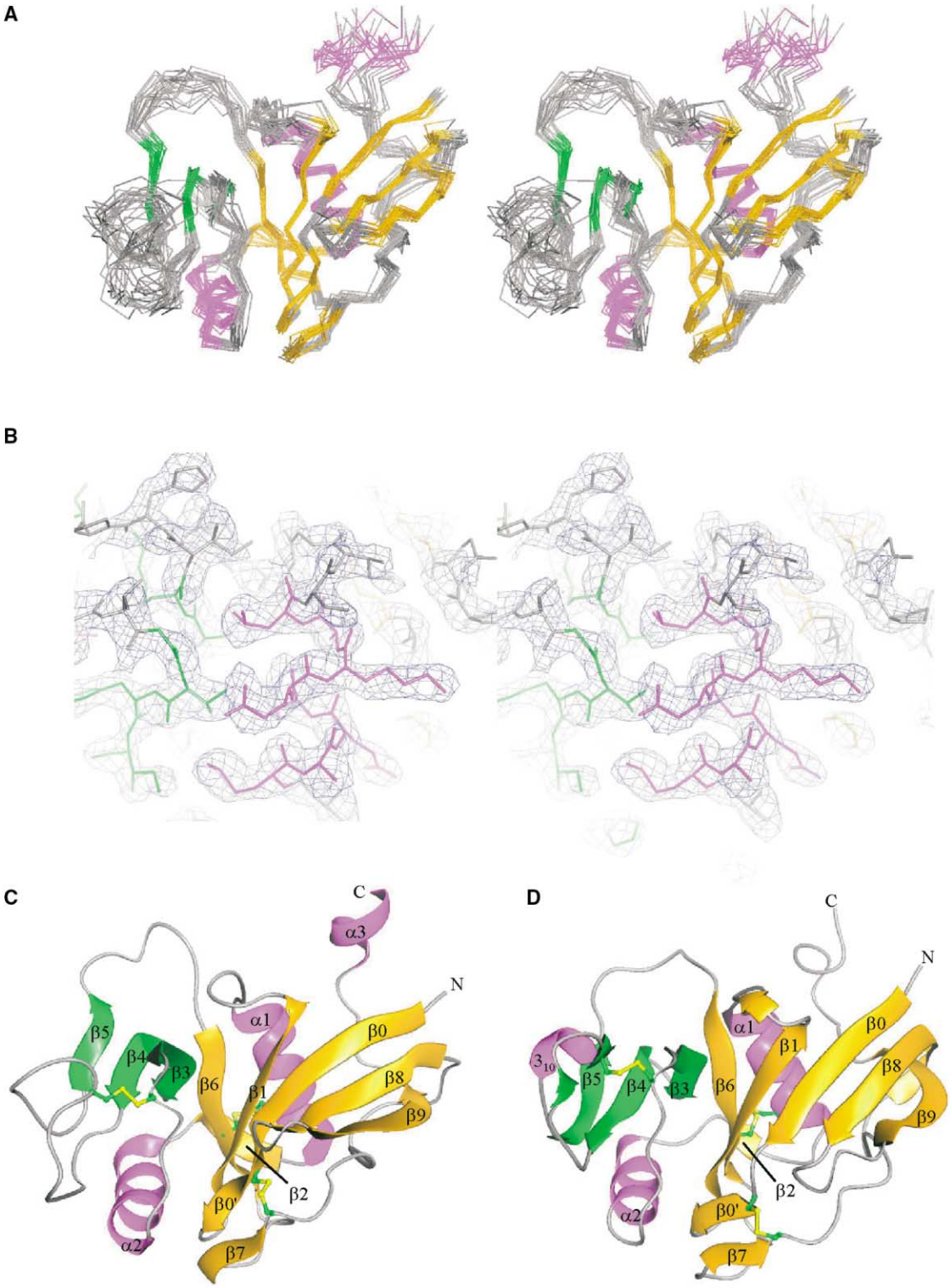

Figure 2. Structures of CD44 HABD Determined by NMR and Crystallography

Residues of strands in sheets I and II are colored gold and green, respectively, and helical residues magenta.

(A) Stereo C^{α} traces for the family of 20 structures determined in solution by NMR spectroscopy superimposed on the backbone heavy atoms in the secondary structure elements.

(B) Representative electron density (shown in stereo) from the averaged SAD phased experimental map, contoured at 0.21 $e^{-}.Å^{-3}$; the final refined model is drawn in cylinder representation.

(C and D) Cartoon representations of the energy minimized averaged structure determined by NMR (C) and one molecule from the final crystallographic structure (D); secondary structure was assigned on the basis of observed NOE contacts, protection of amide protons from solvent exchange (C), and from the coordinates by the DSSP algorithm (D), respectively. Cysteine residues participating in disulfide bonds are drawn in ball and stick representation.

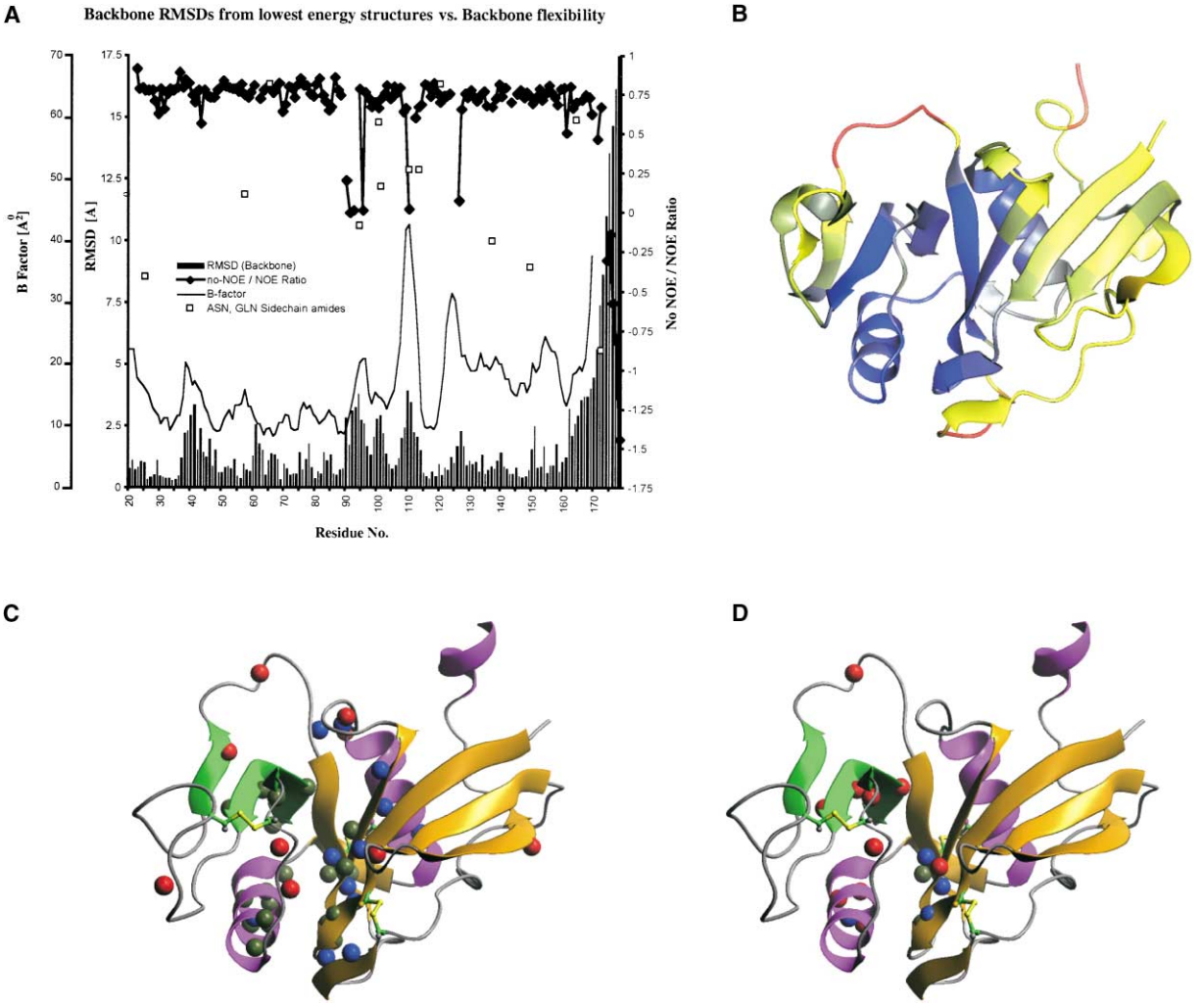

Figure 3. Flexibility and Conformational Change in Models of the CD44 HABD

(A) Quantitative measures of flexibility in the NMR and crystallographic models are plotted as a function of residue number. Parameters plotted are (1) mean main chain temperature factor (thin continuous line), (2) no-NOE/NOE ratio for main chain amides (line with solid diamonds), (3) the no-NOE/NOE ratio for assigned side chain amides (Asn/Gln residues; open squares), and (4) rms coordinate variance among the final family of 20 structures calculated from NMR constraints (histogram).

(B) Cartoon representation of the CD44 fold (determined from crystallography), colored according to C^{α} temperature factor, with saturated blue indicating $B \leq 10 Å^2$ and saturated red indicating $B \geq 40 Å^2$.

(C and D) Slow exchanging amide protons (spheres) mapped onto ribbon representations of the solution structure, detectable after 2.75 (C) and 18.5 (D) hr. Assigned protons detected in both free and HA_6 bound $CD44^{20-178}$ are depicted in gray green. Those observed in only the free or only the bound protein are colored blue or red, respectively.

position to the HA binding surface in TSG-6. An earlier study (Peach et al., 1993) also found that residues in the N- and C-terminal extensions of CD44 (i.e., Arg29, Arg150, Arg154, Lys158, and Arg162) contribute to HA binding. From Figure 5 it can be seen that Arg162 is close to Arg38 but that the others are at more distant positions on this lobe of the HABD and do not form a continuous binding surface (e.g., Arg29, like Lys68, is on the rear face of the domain). The widespread nature of these putative functional residues was unexpected and is difficult to reconcile with a single HA binding configuration within CD44.

Detailed Analysis of the HA Binding Surface by NMR

To further characterize HA binding, we studied the interaction between CD44 and HA oligomers (HA_4–HA_{10}) using NMR. Initial ^1H-^{15}N-HSQC spectra acquired over a range of HA_n to protein stoichiometries (0:1 to 4:1) indicated that the protein/HA complex is in a slow exchange regime on the NMR time scale for all the oligomers tested (see Supplemental Figure S4 for HA_6). Consequently, it was necessary to assign the HABD de novo in its HA-bound form to unambiguously identify residues with perturbed resonances; HA_6 was the shortest oligomer to give essentially full perturbation of the protein. Spectra acquired for $CD44^{20-178}$ in the presence of HA_6 (HA:CD44 ratio 4:1) allowed assignments of more than 72% of the backbone N^H, H^N, and C^{α} nuclei in the complexed HABD.

Comparison of the chemical shifts for the free and bound states of $CD44^{20-178}$ (Supplemental Figure S5) indicates that a large number of nuclei are significantly perturbed upon interaction with HA_6, using the threshold

Table 2. Crystal and Data Collection

Space group	$P2_12_12_1$
Cell dimensions	a = 48.9 b = 77.3 c = 87.6 $\alpha = \beta = \gamma = 90$
Resolution range (Å) (overall, highest shell)	23.31–2.20, 2.32–2.20
R_{sym}[a] (overall, highest shell)	0.076, 0.181
$I/\sigma(I)$ (overall, highest shell)	7.4, 3.1
Anomalous completeness(%) (overall, highest shell)	99.4, 98.2
Multiplicity (overall, highest shell)	3.8, 3.5
Wilson B factor (Å^2)	19.11
Phasing	
Figure of merit[b] (after SOLVE, after RESOLVE)	0.30, 0.59
Refinement	
Number of nonhydrogen protein atoms	2334
Number of water molecules	262
R_{cryst} (R_{free})[c] (23.31–2.20, 2.28–2.20)	0.18 (0.24), 0.19 (0.24)
Rms deviations from ideal bond lengths (Å)	0.011
Rms deviation from ideal bond angles (°)	1.309
Structure quality according to PROCHECK[d]	
Residues in core regions	224
Residues in additional allowed regions	37
Residues in generously allowed regions	1
Residues in disallowed regions	0
REFMAC maximum likelihood estimated rms coordinate error (Å)	0.141
Rms B factor difference for covalently bonded atoms (Å^2) (main chain, side chain)	0.657, 1.824

[a] $R_{sym} = \sum_h \sum_j ||I_{h,j}| - |\bar{I}_h|| / \sum_h \sum_j |I_{h,j}|$ where $I_{h,j}$ is the intensity of the jth observation of unique reflection h.

[b] Mean figure of merit from SOLVE SAD phasing (first value) and after RESOLVE averaging/solvent flattening (value in parentheses).

[c] $R_{cryst} = \sum_h ||Fo_h| - |Fc_h|| / \sum_h |Fo_h|$. All measured reflections were included in refinement except for the 5% disjoint set that was used for calculation of the R_{free}.

[d] Laskowski et al., 1996.

values ($\Delta\delta$ for $H^N \geq 0.2$ ppm, $N^H \geq 1.0$ ppm, $C^\alpha \geq 0.5$ ppm, and $C' \geq 1.0$ ppm) described previously for TSG-6/HA interactions (Blundell et al., 2003). As seen in Figure 5, amino acids that have significant chemical shift changes for all of their H^N, N^H, and C^α nuclei are clustered in three groups (colored yellow). Residues Lys38 and Gly40–Ile44 make up a central patch, with two smaller patches on either side formed from Arg154–Tyr155, and Asn164/Glu166; these segments of the protein experience some of the largest HA-induced shift changes (Supplemental Figure S5). Other amino acids that either have perturbations of at least one H^N, N^H, or C^α nuclei (red on Figure 5; including Arg78, Tyr79, and Arg150) or are perturbed but unassigned (gray; e.g., Arg29 and Arg162), extend these patches. Thus, the residues that are affected by HA binding are mostly confined to the "front" face of the HABD structure, where they overlap appreciably with the ligand binding residues predicted from mutagenesis.

However, certain amino acids identified as functional by mutagenesis do not experience significant backbone shifts on HA_6 binding. With the exception of Lys68, these all lie adjacent to perturbed regions of the structure. For example, Asn100 and Asn101 are likely to be functional since both have backbone shift perturbations close to cutoff levels, and residue 101 in mouse CD44 forms part of the epitope for the HA blocking mAbs KM81 and KM114 (see Figure 6). Lys68 is on the opposite face of the HABD (Figure 5D), in a region that has few chemical shift perturbations approaching threshold levels (residues 60–72, Supplemental Figure S5), and is thus unlikely to interact with HA. In this regard, Arg29, an amino acid with an unassigned chemical shift change, is also on this "back" face of the protein.

In summary, the NMR analyses define a HA-interaction surface on the Link module, centered on Arg41, Tyr42, Arg78, and Tyr79, and extended on either side by Lys38, Asn100, and Asn101, with residues in the C-terminal extension (e.g., Arg150, Arg154, and Arg162) providing additional contacts (see below). However, the distribution of chemical shift changes indicates additional complexity in the binding event (i.e., a conformational rearrangement and/or multiple binding surfaces).

HA Binding Induces a Conformational Change in CD44

The widespread nature of the backbone shift perturbations is in marked contrast to the TSG-6 Link module, where the HA-induced shift changes are all localized around the ligand binding groove (Blundell et al., 2003), one side of which is formed by the $\beta4/\beta5$ loop in which a conformational rearrangement occurs (Kahmann et al., 2000; Blundell et al., 2003).

Therefore, the effect of HA binding on the $CD44^{20\text{-}178}$ structure was investigated further by determining the pattern of slow exchanging amide protons (sH^N) in the HABD/HA_6 complex compared to the free protein. Twenty-one common sH^N were seen in free and bound CD44 at the initial time point (2.75 hr; green in Figure 3C). Interestingly, twelve sH^N were detected only in the free CD44, and nine sH^N were observed only in the complex (colored blue and red, respectively), confirming that HA binding causes a significant structural perturbation. Six of the sH^N protons observed only in free CD44 (residues 34, 37, 38, 118, 120, and 128) are involved in form-

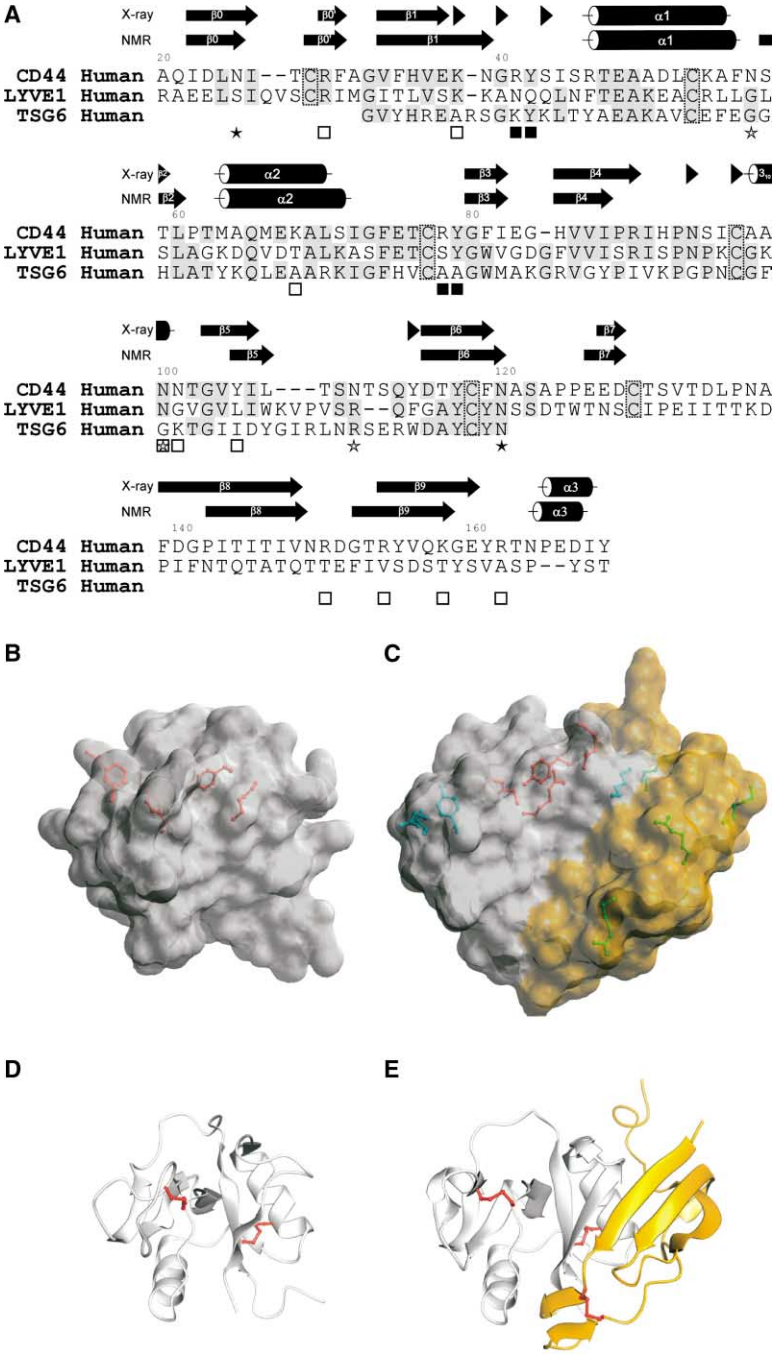

Figure 4. Comparison of the HA Binding Domains from CD44, TSG-6, and LYVE-1

(A) Sequence alignment of the HABDs. Identical residues are shaded gray, and conserved cysteines are indicated by dotted boxes. Secondary structural elements, determined from solution or crystal structures, are shown above the alignment. CD44 residues implicated as "important" or "essential" for HA binding are denoted with open or solid squares, respectively. Sites for N-glycosylation are indicated by open asterisks, which are filled in cases where the glycosylation is believed to modulate HA affinity.

(B and C) Semitransparent molecular surface representations of the TSG-6 Link module (Blundell et al., 2003) (B) and CD44 HABD (C) in the same relative orientations; Link module surfaces are colored gray, and CD44 HABD extensions colored gold. Side chain positions for HA binding residues in CD44 (essential, red; important, light blue [Link module], green [present in extensions]) are compared with those in the TSG-6 Link module (red; see Mahoney et al., 2001b).

(D and E) Cartoon representation of the TSG-6 Link module (D) and CD44 HABD (E) folds in the same orientations as for (B) and (C), respectively; Link modules are colored white, CD44 HABD extensions colored gold, and disulfide bonds colored red (ball and stick representation).

ing H bonds across β strands (β0:β1, β1:β6, and β0:β7; see Supplemental Figure S2). Clearly these bonds are destabilized on interaction with HA. As shown in Figure 3C, most of the sH^N protons seen only in free CD44 (blue) are clustered close to the interface between the Link module and the N- and C-terminal extensions, lying on a central axis. This suggests that the interaction of HA with CD44 induces a change in relative orientation of these subdomains, with a corresponding alteration in the H bond network (see Supplemental Data at http://www.molecule.org/cgi/content/full/13/4/483/DC1 for further details). It should be noted that all but one of the sH^N assigned in the free protein were found to be involved in interresidual H bonds interactions in the crystal structure (Supplemental Figure S6), indicating that this is representative of the CD44 HABD in its unbound state.

The nine sH^N observed solely in the HABD/HA_6 complex are mostly in loop regions that may become more structured and less dynamic on HA binding (red on Figure 3C); only the amide proton of Leu107 is involved in an interstrand H bond (see Supplemental Figure S2). The finding that almost twice as many sH^N are detected for the HA_6 bound protein compared to free CD44 after 18.5 hr indicates a ligand-induced stabilization of the HABD structure (see Figure 3D). The eleven long-lived

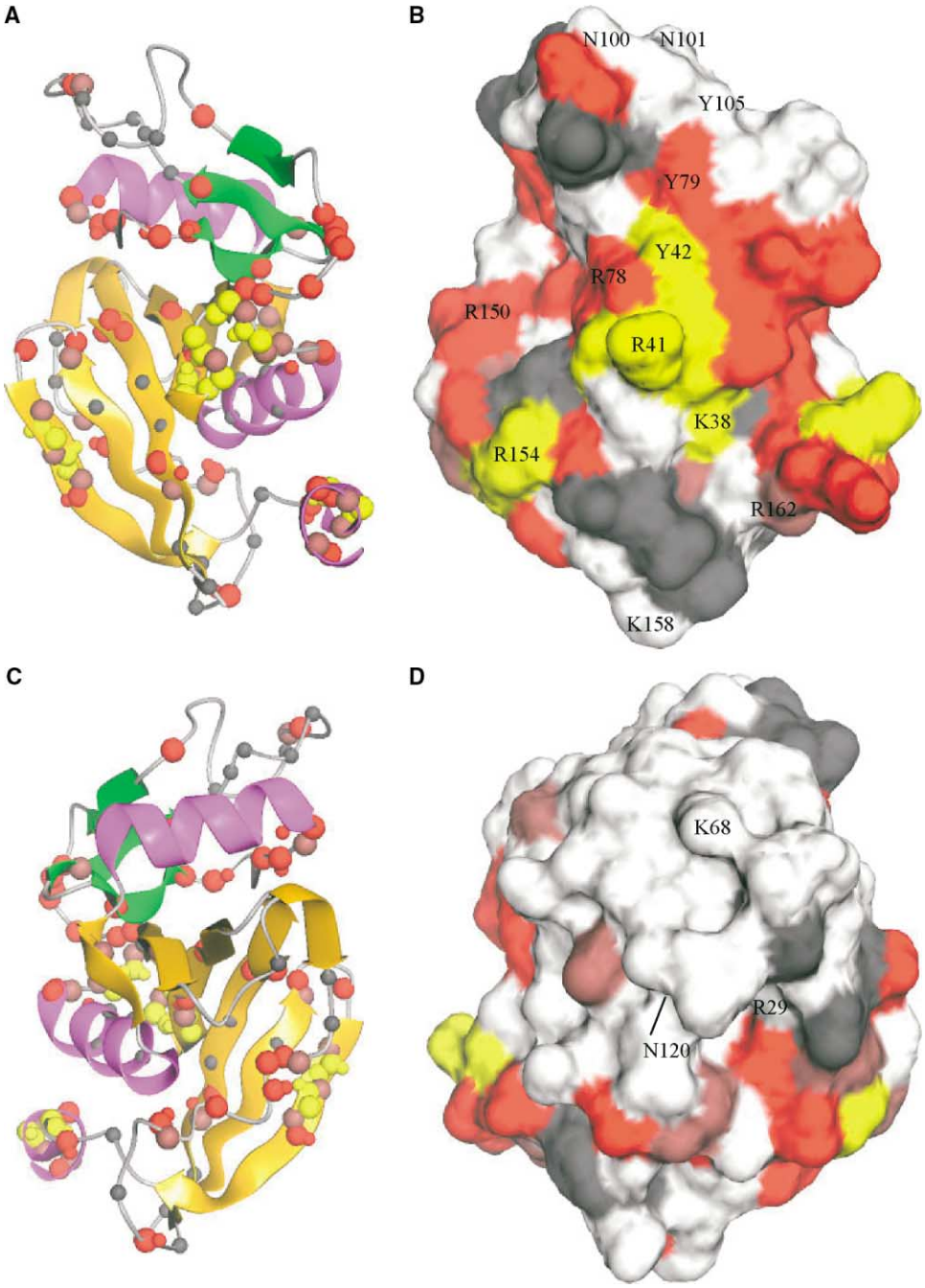

Figure 5. Mapping the CD44 HA Binding Surface

The HABD (X-ray structure) is shown as ribbon (A and C) or contact surface (B and D) representations; structures in (A) and (B) are rotated 180° around the vertical axis relative to (C) and (D). Nuclei (A and C), shown as spheres, or their related amino acids (B and D) are colored according to degree of chemical shift perturbations on binding HA_6. Yellow denotes residues with H^N, N^H, and C^α nuclei that all exhibit a significant perturbation (see text). Red indicates amino acids that have at least one affected H^N, N^H, or C^α nucleus. Residues with perturbed carbonyl (C′) resonances are colored dark pink; 93% and 71% of the assignable backbone C′ were derived for the free and bound protein, respectively. C^α nuclei of residues where H^N, N^H, and C^α were unassigned in the complex are shown in gray (see Supplemental Figure S5). Positions of residues implicated in HA binding on the basis of site-directed mutagenesis are shown in (B) and (D) for comparison.

sH^N observed only in the complex (red) are all found within the Link module, being mostly located on the α2 helix, β3, and β4 strands. Overall, these data suggest that the interaction of HA with CD44 stabilizes the Link module portion of the HABD while inducing a change in its orientation relative to the rest of the structure.

Determining the Size of HA Accommodated in the Binding Site

Comparison of ^{1}H-^{15}N-HSQC experiments on $CD44^{20-178}$ in the presence of HA_4, HA_6, HA_8, or HA_{10} revealed that these oligomers all give rise to very similar shift perturbations (see Supplemental Figure S7). This indicates that

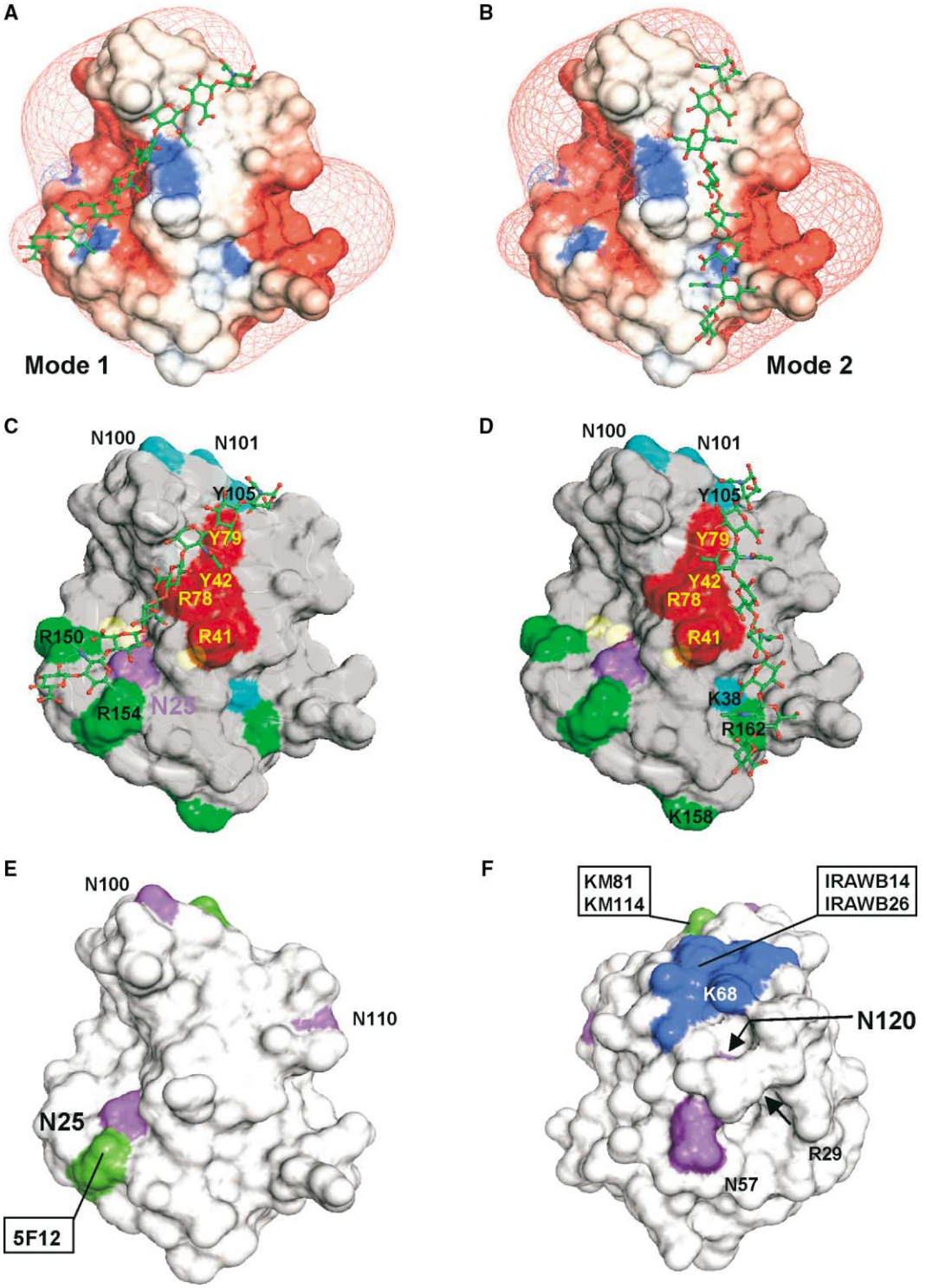

Figure 6. Hypothetical Models for HA Binding to the CD44 HABD

(A–D) HA octasaccharides are shown in two different orientations (i.e., mode 1 [A and C] or mode 2 [B and D]) on a molecular surface representation of CD44. In (A) and (B) the surface is colored according to electrostatic potential (GRASP program), with red and blue corresponding to potentials ≤ -2 kT/e and ≥ 2 kT/e, respectively; the map is contoured at a level of -2 kT/e. In (C) and (D), residues implicated in HA binding are color-coded as in Figure 4. Likely contact residues for each binding mode are appropriately annotated. Yellow spheres visible through the semitransparent surface denote positions of nuclei that show significantly perturbed chemical shifts upon binding to HA_6, but not with HA_4. In (E) and (F) positions of N-linked glycosylation sites (magenta) and epitopes of monoclonal antibodies (activating, blue; blocking, green) are shown on the front (E) and back (F) faces of the HABD; (E) is in the same orientation as (A)–(D). On (F) the locations of residues Arg29 and K68, the mutation of which reduces HA binding, are also indicated; note that the view of Asn120 is largely occluded at this viewing angle by the loop between β strands 6 and 7.

they bind at essentially the same site on CD44 and all cause a similar ligand-induced conformational change. However, there are several distinct differences between the effects of HA_4 and the longer oligomers (see Supplemental Data), and a single clear (but unassigned) difference in the $CD44^{20-178}$ HSQC spectra in the presence of HA_6 and HA_8. This suggests that an 8-mer is likely to be the minimum size of HA that completely occupies the HA binding site in the nonglycosylated CD44 HABD used here, consistent with the results of the ELISA experiments described above.

Does CD44 Have Two Different Modes of HA Binding?

The residues within CD44 strongly implicated in HA binding (Lys38, Arg41, Tyr42, Arg78, Tyr79, Asn100, Asn101, Arg150, Arg154, Arg162; based on mutagenesis and NMR shift data) map onto a single face of the HABD; this region is surrounded on all sides by a shell of negative electrostatic charge that may help to guide the polyanionic HA toward its docking site via charge repulsion (Figure 6 and Supplemental Figure S8). However, these residues are too widespread to be able to interact with a single octasaccharide. As shown in Figure 6, this could be explained if CD44 was able to accommodate HA molecules in two different binding positions or modes. In the first mode, HA_8 would thread from Asn100/Asn101 at one end toward Arg150/Arg154 at the other (Figures 6A and 6C). This arrangement is consistent with the different backbone amide perturbations observed for Thr27, Cys28, and Lys38 (yellow on Figure 6C) in the presence of HA_4 and HA_6 (see Supplemental Data) and with the fact that Arg 154 forms part of the epitope for the HA-blocking mAb 5F12 (Liao et al., 1995). Interestingly, the HA in mode 1 is also in a similar orientation to an octasaccharide modeled onto the bound conformation of the TSG-6 Link module (Blundell et al., 2003). Furthermore, recently reported cross-saturation experiments (Takeda et al., 2003), when reinterpreted on the basis of our 3D structure, are consistent with this mode of interaction (it should be noted that neither Arg41 nor Tyr42 were identified as HA binding residues in the Takeda study).

In the alternative mode 2, the HA molecule would make bonded interactions with Lys38 or Arg162 and lie more completely in the uncharged region of the interaction surface. Since modes 1 and 2 overlap (Figures 6C and 6D), it would not be possible for them to be used simultaneously in an individual CD44 molecule. It is possible that changes in glycosylation in vivo (e.g., at Asn25) could determine which mode of interaction is utilized by CD44.

Regulation of CD44 Function by N-Glycosylation

It is well established in vitro that inhibition of N-glycan biosynthesis or enzymatic hydrolysis of N-linked glycans or their terminal sialic acid residues can "activate" CD44 in certain cell types (Katoh et al., 1995, 1999; Lesley et al., 1995). Of the five N-glycosylation sites within the CD44 HABD, two in particular, Asn25 and Asn120, have been identified as important on the basis that their individual mutation to serine leads to binding activation (English et al., 1998). Moreover, both were shown to carry terminal α2,3 linked sialic acid whose cleavage by neuraminidase led to activation of HA binding (English et al., 1998).

Although it is readily apparent from our studies how a negatively charged sugar chain attached to Asn25 could directly obstruct HA binding (see Figure 6), this is less obvious for Asn120, which is on the rear face of the HABD, close to Arg29 (Figure 6F). A likely explanation is that charged glycans attached at Asn120 interfere with CD44 self-association, a process believed to have physiological importance in regulating HA binding avidity (see Lesley et al., 2000). Such clustering can also be induced by "superagonist" mAbs such as IRAWB14 and IRAWB26 that trigger large increases (50- to 100-fold) in HA binding (Lesley et al., 1993; Zheng et al., 1995), possibly by orientating the crosslinked CD44 molecules for optimal binding to adjacent sites along the same HA molecule (Lesley et al., 2000). Interestingly, our studies reveal that the shared binding site for IRAWB14 and IRAWB26 (D^{64}QMKLALS^{71} in mouse CD44, corresponding to A^{64}QMEKALS^{71} in the human [Zheng et al., 1995]) lies immediately adjacent to Asn120 in the CD44 structure and is centered on Lys68, one of only two residues on the rear face of the HABD whose mutation influences HA binding (Bajorath et al., 1998).

These data are therefore consistent with the notion that the rear face of CD44 may be involved in regulating HA binding, either by enhancing the tandem juxtaposition of monomer HABDs in a fashion analogous to IRAWB14/IRAWB26 crosslinking or by effecting a conformational change leading to high-affinity binding. Specifically, the presence of terminal sialic acid residues on Asn120 could prevent formation of CD44 dimers or higher oligomers through charge repulsion. While confirmation of such a role for Asn120-linked glycans must await experimental validation, it is worth noting that sialylated O-glycans have been proposed to regulate receptor self-association in the case of the T cell coreceptor CD8 and the leukocyte receptor tyrosine phosphatase CD45 (Moody et al., 2001; Xu and Weiss, 2002). Such a role for N-glycan chains, as exemplified by CD44, is to our knowledge unique.

In conclusion, the enlarged CD44 HABD allows for regulation of ligand binding in response to proinflammatory cytokines and other extracellular factors that transduce changes in receptor N-glycosylation. The sequences flanking the Link module, which are unique to CD44, play a pivotal role by providing the attachment sites for key N-glycans (Asn25 and Asn120) whose appropriate modification may either directly block HA binding or reduce receptor avidity by preventing critical homooligomerization. This regulation of binding is fundamental to many of the biological roles in which CD44 participates, such as transendothelial migration of leukocytes from the circulation to the sites of inflammation and the invasion of end organ tissue by metastatic tumor cells (Bartolazzi et al., 1994; De Grendele et al., 1997; Mikecz et al., 1995; Katoh et al., 2003). The structure presented in this paper provides a platform for further studies leading to the rational design of small molecule inhibitors of HA binding and CD44 self-association as potential therapeutic agents for the treatment of inflammatory and malignant diseases.

Experimental Procedures

Reagents

The monoclonal antibodies BRIC-235 and F10.44.2 were obtained from Professor D. Anstee (Bristol, UK) and SouthernBiotech (Birmingham, AL) respectively. Selenomethionine was purchased from Calbiochem (Nottingham, UK). ^{15}N ammonium chloride (99.9 atom %) and D-glucose-$^{13}C_6$ (99.2 atom %) were supplied by CK Gas Products (Cambridge, UK). Soluble glycosylated CD44Fc was produced as described in Banerji et al. (1998). HA oligosaccharides of defined length (HA_4, HA_6, HA_8, HA_{10}, HA_{12}, and HA_{20-24}) were prepared either as described in Lesley et al. (2000) or Mahoney et al. (2001a).

Cloning and Expression of the Soluble CD44 HA Binding Domain

The ectodomain construct $CD44^{20-178}$ comprising the final residue from the signal sequence and the first 158 residues from the N terminus of the mature human protein was prepared by PCR amplification from a full-length cDNA clone using the forward primer CCAT GGccatgg$_{56}$CGCAGATCGATTTGAATATA containing an NcoI site (lowercase) and the reverse primer TGGAagatct*TCA*C$_{530}$A CGTCAT CATCAGTAGGG containing a BglII site for ligation into the expression vector pET19b. Protein was expressed and purified to homogeneity as described previously (Banerji et al., 1998). For heavy metal replacement the construct was transformed into *E. coli* B834 DE3 and grown at 30°C in M9 minimal media supplemented with Mg_2SO_4 (mM), $FeSO_4$ (25 μg/ml), glucose (0.4% [w/v]), amino acids (40 μg/ml each, with seleno-L-methionine substituted for L-methionine) and vitamins (riboflavin, niacinamide, pyridoxine monohydrochloride, thiamine [1 μg/ml each]). For single (^{15}N-) and double (^{13}C,^{15}N-) labeling, bacteria were grown at 30°C in enriched M9 media (0.6% [w/v] anhydrous Na_2HPO_4, 0.3% [w/v] KH_2PO_4, 0.05% [w/v] NaCl, 0.085% [w/v] yeast nitrogen base, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$) supplemented with ^{15}N H_4Cl (0.052% [w/v]) and $^{13}C_6$ glucose (0.15% [w/v]), as appropriate.

The molecular masses of $CD44^{20-178}$ and seleno-$CD44^{20-178}$ assessed by electrospray ionization mass spectrometry were 17579.5 and 17673.5 Da, respectively, consistent with the expected sizes of the polypeptides after cleavage of the N-terminal methionine (e.g., 17,580.4 Da for unmodified protein). Mass spectrometry indicated that there was 100% and 98% incorporation of ^{15}N and ^{13}C isotopes, respectively.

Crystallization of $CD44^{20-178}$, Data Collection, and Analysis

$CD44^{20-178}$ was crystallized by hanging-drop vapor diffusion. Droplets of 1 μl protein sample +1 μl well solution were suspended above well solutions of 12% PEG 3350, 50 mM NaCl, and 10% glycerol. Crystals grew to $5 \times 5 \times 300$ μm^3 over 1 week. The same conditions were used for crystallization of the selenomethionine protein except that monothioglycerol (100 μg/ml final concentration) was added to both the protein and well solutions.

SAD data were collected from the SeMet $CD44^{20-178}$ crystal to 2.2 Å at beamline ID14.4 of the ESRF, using 0.9795 Å wavelength radiation. A data set from $CD44^{20-178}$ crystal was also collected at Elettra at 1.5 Å wavelength to enhance the anomalous signal of sulfur (absorption edge at 2475 eV [Weiss et al., 2001]). For this data set, a helium path was employed. SOLVE was able to determine positions of selenium atoms in the ESRF data set and to produce electron density of reasonable, but not readily interpretable, quality. Determination of the positions of the sulfurs in $CD44^{20-178}$ was accomplished by crossphasing an anomalous difference Fourier of the Elettra dataset. These sites defined the noncrystallographic symmetry (NCS) transformation relating the two monomers in the asymmetric unit, permitting structure determination by the programs SOLVE/RESOLVE (Terwilliger and Berendzen, 1999), O (Jones et al., 1991), and programs of the CCP4 suite (CCP4, 1994). The initial maps, into which RESOLVE was able to automatically build a 70% complete structure, were calculated at 2.6 Å resolution. This model could readily be completed in the first round of interactive model building in O, and was refined with tight NCS restraints in the program REFMAC using all data to 2.2 Å resolution. Water positions were determined manually by inspection of the chemical environment of high ($>3\sigma$) peaks in difference Fouriers. Ten rounds of REFMAC and interactive refinement were performed. At round eight, the program PROCHECK (Laskowski et al., 1996) was used to check the model. This revealed no residues in disallowed regions of the Ramachandran plot, and 3 of the 304 visible residues in generously allowed regions. Of these, two residues (residues 107 and 112 of chain A) were in weak electron density, and their conformations were reset so as to better agree with the database of known structures. The last, residue 126, was clearly defined in electron density and presumably lies in a region of slight local strain. In the last two rounds of refinement, the tight NCS restraints were released. This decreased R_{cryst} by 0.03 and R_{free} by 0.01. Data collection and refinement statistics are shown in Table 2. The structure appears under Protein Data Bank accession code 1UUH.

NMR Spectroscopy

NMR samples were prepared from lyophilized $CD44^{20-178}$ in 10% (v/v) or 100% (v/v) D_2O (containing 0.02% [w/v] Na-azide) and adjusted to pH 6.5. All NMR experiments were recorded at 25°C at proton frequencies of 500, 600, or 750 MHz. Multidimensional data sets were recorded in a phase-sensitive manner with quadrature detection by the States/TPPI method (States et al., 1982). The ^1H carrier frequency was set to the water resonance (4.75 ppm) and broadband decoupling of the heteronuclei was achieved using GARP or WALTZ-16 sequences at field strengths of 833 Hz (^{15}N) and 2500 Hz (^{13}C) (Shaka et al., 1985).

Assignment (backbone and side chain) was based on HNCA, CBCA(CO)NH, CC(CO)NH, ^1H-^{15}N-TOCSY-HSQC (mixing time 25 ms), ^1H-^{15}N-NOESY-HSQC (mixing times 70 or 150 ms), and HCCH-TOCSY (mixing time 11 ms) spectra recorded on 1 mM protein samples at 600 or 750 MHz. The $^3J_{HN\text{-}H\alpha}$ coupling constants were determined using line width analysis of 2D-NOESY and ^1H-^{15}N-HMQC-J experiments. NOE correlations were derived from ^{15}N-edited and 2D-D_2O-NOESY experiments (mixing times of 70 ms). Amide exchange with solvent was observed with ^{15}N-^1H-HSQC experiments recorded on a 0.3 mM sample (500 MHz) at 105 min intervals (following an initial delay of 60 min) after the addition of 5 mM MES (pH 6.5) in 100% D_2O to lyophilized protein; this was also performed in the presence of 1.2 mM HA_6. The ^{15}N-{^1H} NOE were measured at 500 MHz by comparison of spectra recorded either with (NOE) or without (no-NOE) ^1H saturation, as described in Pickford et al. (2001). Data were processed using FELIX 2.3 (Biosym Inc.), and referenced and analyzed with XEasy as described in Kahmann et al. (2000). Peak heights in the ^{15}N-{^1H} NOE experiment were measured with the program SPARKY (Goddard and Knelle, 2003). NOE intensities for each data set were measured and calibrated using interproton distances in regions of regular secondary structure, and converted into four distance restraint categories with upper limits of 2.7 Å, 3.5 Å, 5.0 Å, and 6.2 Å. These restraints were used within the program CNS v1.1 (2001; see Brunger et al., 1998) in an ab initio simulated annealing protocol which utilizes the SOPHIE algorithm (Pickford et al., 2001), with $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shift values used to provide additional structural information (see Wishart and Sykes, 1994). Hydrogen bond restraints were included as described in Pickford et al. (2001). A total of 500 structures were calculated, and the 30 of lowest target function refined by three further cycles of the cooling step from the ab initio calculation, followed by extensive restrained energy minimization (see Pickford et al., 2001). The 20 lowest energy structures were used to calculate a mean structure using CNS (Protein Data Bank accession code 1POZ).

^1H-^{15}N-HSQC experiments were performed on 0.3 mM ^{15}N-labeled $CD44^{20-178}$ in the presence of HA_4, HA_6, HA_8, or HA_{10} at sugar to protein stoichiometries of 0:1, 0.5:1, 1:1, 2:1, and 4:1. HNCA, HN(CO)CA, and HNNH (as in Panchal et al., 2001), but modified to be a 2D experiment (S.A. Colebrooke, G. Verdone, and I.D.C., unpublished data); spectra were acquired for 1 mM $CD44^{20-178}$ samples containing HA_6 at a 4-fold molar ratio. Spectra from free and bound protein were used to determine $^1H^N$ $^{15}N^H$ and $^{13}C^{\alpha}$ assignments for the complex. HNCO spectra were recorded on both free and HA_6-containing samples for measurement of $^{13}C'$.

HA Binding Assays

For the standard assay, 96-well microtiter plates (Nunc Maxisorp) were coated (100 μl/well, 6 hr, 20°C) with samples of $CD44^{20-178}$ or

CD44Fc (5–40 μg/ml) in 15 mM sodium carbonate buffer (pH 9.3) followed by washing in PBS and blockage of unreacted sites in PBS containing 1% (w/v) BSA, 0.05% (w/v) Tween 20. Plates were then incubated with biotinylated HA (5 μg/ml, 20°C, 1 hr), prepared as described previously (Mahoney et al., 2001b), and bound HA detected with horseradish peroxidase-conjugated streptavidin (Dako) and *O*-phenylenediamine on a Bio-Rad microplate reader at OD_{490nm}.

The relative affinity of CD44 for other GAGs was determined in a modification of the standard assay in which $CD44^{20\text{-}178}$ was coated at 20 μg/ml, and competing GAGs (C4S and C6S; Seikagaku Corporation) were added to the plates (10 to 1000 μg/ml) 15 min prior to the addition of biotinylated HA. A similar modification was used to assess competition with defined HA oligosaccharides (HA_4 to $HA_{22\text{-}24}$).

Acknowledgments

We thank Robin Aplin for performing mass spectrometry. We are grateful to the beamline scientists at the synchrotrons at ESRF (Grenoble) Elettra (Trieste) and members of the cell-cycle group at the LMB, Oxford for help with data collection. This work was generously funded by the Medical Research Council (D.G.J. and A.J.D.) and the Arthritis Research Campaign (grants D0525, D0540, D0569, and M0625). We also acknowledge the financial support for P.T. from Aventis Pharmaceuticals as part of an MRC Collaborative Research Initiate Studentship. NMR spectroscopy was carried out at the Oxford Centre for Molecular Sciences, funded by the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, and the Medical Research Council.

Received: September 3, 2003
Revised: December 24, 2003
Accepted: December 30, 2003
Published: February 26, 2004

References

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

Bajorath, J., Greenfield, B., Munro, S.B., Day, A.J., and Aruffo, A. (1998). Identification of CD44 residues important for hyaluronan binding and delineation of the binding site. J. Biol. Chem. *273*, 338–343.

Banerji, S., Day, A.J., Kahmann, J.D., and Jackson, D.G. (1998). Characterization of a functional hyaluronan-binding domain from the human CD44 molecule expressed in Escherichia coli. Protein Expr. Purif. *14*, 371–381.

Banerji, S., Ni, J., Wang, S.-X., Clasper, S., Su, J., Tammi, R., Jones, M., and Jackson, D.G. (1999). LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph-specific receptor for hyaluronan. J. Cell Biol. *144*, 789–801.

Bartolazzi, A., Peach, R., Aruffo, A., and Stamenkovic, I. (1994). Interaction between CD44 and hyaluronate is directly implicated in the regulation of tumor development. J. Exp. Med. *180*, 53–66.

Blundell, C.D., Mahoney, D.J., Almond, A., DeAngelis, P.L., Kahmann, J.D., Teriete, P., Pickford, A.R., Campbell, I.D., and Day, A.J. (2003). The Link module from ovulation-associated protein TSG-6 changes conformation on hyaluronan binding. J. Biol. Chem. *278*, 49261–49270.

Brunger, 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., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

CCP4 (Collaborative Computational Project, Number 4). (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

Day, A.J., and Prestwich, G.D. (2002). Hyaluronan-binding proteins: tying up the giant. J. Biol. Chem. *277*, 4585–4588.

De Grendele, H.C., Estess, P., and Siegelman, M.H. (1997). Requirement for CD44 in activated T cell extravasation into an inflammatory site. Science *278*, 672–675.

English, N.M., Lesley, J.F., and Hyman, R. (1998). Site-specific de-N-glycosylation of CD44 can activate hyaluronan binding, and CD44 activation states show distinct threshold densities for hyaluronan binding. Cancer Res. *58*, 3736–3742.

Gee, K., Kozlowski, M., and Kumar, A. (2003). TNFα induces functionally active hyaluronan-adhesive CD44 by activating silidase through p38 mitogen-activated protein kinase in lipopolysaccharide-stimulated human monocytic cells. J. Biol. Chem. *278*, 37275–37287.

Goddard, T.D., and Kneller, D.G. (2003). SPARKY (University of California, San Francisco), www.cgl.ucsf.edu/home/sparky/.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Jackson, D.G. (2003). The lymphatics revisited: new perspectives from the hyaluronan receptor LYVE-1. Trends Cardiovasc. Med. *13*, 1–7.

Jackson, D.G., Prevo, R., Clasper, S., and Banerji, S. (2001). LYVE-1, the lymphatic system and tumor angiogenesis. Trends Immunol. *22*, 317–321.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kahmann, J.D., O'Brien, R., Werner, J.M., Heinegard, D., Ladbury, J.E., Campbell, I.D., and Day, A.J. (2000). Localization and characterization of the hyaluronan-binding site on the Link module from human TSG-6. Structure *8*, 763–774.

Katoh, S., Zheng, Z., Oritani, K., Shimozato, T., and Kincade, P. (1995). Glycosylation of CD44 negatively regulates its recognition of hyaluronan. J. Exp. Med. *182*, 419–429.

Katoh, S., Miyagi, T., Taniguchi, H., Matsubara, Y., Kadota, J., Tominaga, A., Kincade, P.W., Matsukura, S., and Kohno, S. (1999). An inducible sialidase regulates the hyaluronic acid binding ability of CD44-bearing human monocytes. J. Immunol. *162*, 5058–5061.

Katoh, S., Matsumoto, N., Kawakita, K., Tominaga, A., Kincade, P., and Matsukura, M. (2003). A role for CD44 in an antigen-induced murine model of pulmonary eosinophilia. J. Clin. Invest. *111*, 1563–1570.

Kincade, P.W., Zheng, Z., Katoh, S., and Hanson, L. (1997). The importance of cellular environment to function of the CD44 matrix receptor. Curr. Opin. Cell Biol. *9*, 635–642.

Kohda, D., Morton, C.J., Parkar, A.A., Hatanaka, H., Inagaki, F.M., Campbell, I.D., and Day, A.J. (1996). Solution structure of the link module: a hyaluronan binding domain involved in extracellular matrix stability and cell migration. Cell *86*, 767–775.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Lesley, J., Kincade, P.W., and Hyman, R. (1993). Antibody-induced activation of the hyaluronan receptor function of CD44 requires multivalent binding by antibody. Eur. J. Immunol. *23*, 1902–1909.

Lesley, J., English, N., Perschl, A., Gregoroff, J., and Hyman, R. (1995). Variant cell lines selected for alterations in the function of the hyaluronan receptor CD44 show differences in glycosylation. J. Exp. Med. *182*, 431–437.

Lesley, J., Hyman, R., English, N., Catterall, J.B., and Turner, G.A. (1997). CD44 in inflammation and metastasis. Glycoconj. J. *14*, 611–622.

Lesley, J., Hascall, V., Tammi, M., and Hyman, R. (2000). Hyaluronan binding by cell surface CD44. J. Biol. Chem. *275*, 26967–26975.

Levesque, M.C., and Haynes, B.F. (1999). TNFα and IL-4 regulation of hyaluronan binding to monocyte CD44 involves posttranslational modification of CD44. Cell. Immunol. *193*, 209–218.

Liao, H.-X., Lee, D.M. Levesque, M.C. and Haynes, B, F. (1995). N-terminal and central regions of the human CD44 extracellular

domain participate in cell surface hyaluronan binding. J. Immunol. *155*, 3938–3945.

Mahoney, D.J., Aplin, R.T., Calabro, A., Hascall, V.C., and Day, A.J. (2001a). Novel methods for the preparation and characterization of hyaluronan oligosaccharides of defined length. Glycobiology *11*, 1025–1033.

Mahoney, D.J., Blundell, C., and Day, A.J. (2001b). Mapping the hyaluronan binding site on the Link module from human tumour necrosis factor-stimulated gene-6 by site-directed mutagenesis. J. Biol. Chem. *276*, 27764–22771.

Mikecz, K., Brennan, F.R., Kim, J.H., and Glant, T.T. (1995). Anti-CD44 treatment abrogates tissue oedema and leukocyte infiltration in murine arthritis. Nat. Med. *1*, 558–563.

Mohamadzadeh, M., DeGrendele, H., Arizpe, H., Estess, P., and Siegelman, M. (1998). Proinflammatory stimuli regulate endothelial hyaluronan expression and CD44/HA-dependent primary adhesion. J. Clin. Invest. *101*, 97–108.

Moody, A.M., Chui, D., Reche, P.A., Priatel, J.J., Marth, J.D., and Reinherz, E.L. (2001). Developmentally regulated glycosylation of the CD8$\alpha\beta$ coreceptor stalk modulates ligand binding. Cell *107*, 501–512.

Noble, P.W. (2002). Hyaluronan and its catabolic products in tissue injury and repair. Matrix Biol. *21*, 25–29.

Panchal, S.C., Bhavesh, N.A., and Hosur, R.V. (2001). Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and 15N sequential correlations in (13C, 15N) labeled proteins: application to unfolded proteins. J. Biomol. NMR *2*, 135–147.

Peach, R.J., Hollenbaugh, D., Stamenkovic, I., and Aruffo, A. (1993). Identification of hyaluronic acid binding sites in the extracellular domain of CD44. J. Cell Biol. *122*, 257–264.

Pickford, A.R., Smith, S.P., Staunton, D., Boyd, J., and Campbell, I.D. (2001). The hairpin structure of the (6)F1(1)F2(2)F2 fragment from human fibronectin enhances gelatin binding. EMBO J. *20*, 1519–1529.

Pure, E., and Cuff, C.A. (2001). A crucial role for CD44 in inflammation. Trends Mol. Med. *7*, 213–221.

Shaka, A.J., Barker, P.B., and Freeman, R. (1985). Computer optimized decoupling scheme for wideband applications and low-level operation. J. Magn. Reson. *64*, 547–552.

States, D.J., Haberkorn, R.A., and Ruben, D.J. (1982). A 2-dimensional nuclear Overhauser experiment with pure phase in 4 quadrants. J. Magn. Reson. *28*, 286–292.

Takeda, M., Terasawa, H., Sakakura, M., Yamaguchi, Y., Kajiwara, M., Kawashima, H., Miyasaka, M., and Shimada, I. (2003). Hyaluronan recognition mode of CD44 revealed by cross-saturation and chemical shift perturbation experiments. J. Biol. Chem. *278*, 43550–43555.

Tammi, M.I., Day, A.J., and Turley, E.A. (2002). Hyaluronan and homeostasis: a balancing act. J. Biol. Chem. *277*, 4581–4584.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.

Toole, B.P., Wight, T.N., and Tammi, M.I. (2002). Hyaluronan-cell interactions in cancer and vascular disease. J. Biol. Chem. *277*, 4593–4596.

Weiss, L., Slavin, S., Reich, S., Cohen, P., Shuster, S., Stern, R., Kaganovsky, E., Okon, E., Rubinstein, A.M., and Naor, D. (2000). Induction of resistance to diabetes in non-obese diabetic mice by targeting CD44 with a specific monoclonal antibody. Proc. Natl. Acad. Sci. USA *97*, 285–290.

Weiss, M.S., Sicker, T., Djinovic-Carugo, K., and Hilgenfeld, R. (2001). On the routine use of soft X-rays in macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *57*, 689–695.

Wishart, D.S., and Sykes, B.D. (1994). The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J. Biomol. NMR *4*, 171–180.

Xu, Z., and Weiss, A. (2002). Negative regulation of CD45 by differential homodimerization of the alternatively spliced isoforms. Nat. Immunol. *3*, 764–771.

Zheng, Z., Katoh, S., He, Q., Oritani, K., Miyake, K., Lesley, J., Hyman, R., Hamik, A., Parkhouse, R.M.E., Farr, A.G., and Kincade, P.W. (1995). Monoclonal antibodies to CD44 and their influence on hyaluronan recognition. J. Cell Biol. *130*, 485–495.

Accession Numbers

The crystal structure appears under Protein Data Bank accession code 1UUH, and the NMR structure under 1POZ. Chemical shift assignments have been deposited at BioMagResBank (accession code 6093).