Published in final edited form as:

Biochemistry. 2016 January 19; 55(2): 262–276. doi:10.1021/acs.biochem.5b01148.

NMR insight into the multiple glycosaminoglycan binding modes of the Link module from human TSG-6

Younghee Park¹, Thomas A. Jowitt², Anthony J. Day^{2,*}, and James H. Prestegard^{1,*}

¹Complex Carbohydrate Research Center, 315 Riverbend Road, University of Georgia, Athens, GA 30602, USA

²Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, University of Manchester, Oxford Road, Manchester M13 9PT, UK

Abstract

Tumor necrosis factor-stimulated gene-6 (TSG-6) is a hyaluronan (HA) binding protein that is essential for stabilizing and remodelling the extracellular matrix (ECM) during ovulation and inflammatory disease processes such as arthritis. The Link module, one of the domains of TSG-6, is responsible for binding hyaluronan and other glycosaminoglycans (GAGs) found in the ECM. In this study, we used a well-defined chondroitin sulfate (CS) hexasaccharide (C444S) to determine the structure of the Link module, in solution, in its chondroitin sulfate bound state. A variety of NMR techniques were employed, including chemical shift perturbation, residual dipolar couplings (RDCs), NOEs, spin relaxation measurements, and paramagnetic relaxation enhancements (PREs) from a spin-labeled analog of C444S. The binding site for C444S on the Link module overlapped with that of HA. Surprisingly, C444S binding induced dimerization of the Link module (as confirmed by analytical ultracentrifugation), and a second weak binding site that partially overlapped with a previously identified heparin site was detected. A dimer model was generated using chemical shift perturbations and RDCs as restraints in the docking program HADDOCK. We postulate that the molecular cross-linking enhanced by the multiple binding modes of the Link module may be critical for remodeling the ECM during inflammation/ovulation and may contribute to other functions of TSG-6.

Graphical Abstract

*Correspondence: J. H. Prestegard: Phone: (706) 542-6281. Fax: (706) 542-4412. jpresteg@ccrc.uga.edu. A. J. Day: Phone: 44 (0)161 27 51495. Fax: 44 (0)161 27 55082. anthony.day@manchester.ac.uk.

SUPPLEMENTARY DATA

The following supporting material may be accessed free of charge online at <http://pubs.acs.org>: ¹³Cα and ¹³Cβ chemical shift perturbation figures for C444S binding along with ¹⁵N-¹H chemical shift perturbations for on binding two additional ligands,

C664S and C664S. Titration profiles fit to binding constants and exchange rates for the latter two. Analytical ultracentrifugation data, fits and residuals. A figure depicting binding of CUB modules to the Link_TSG6 dimer.

Keywords

Link module; Dimerization; TSG-6; Chondroitin sulfate; NMR spectroscopy; Chemical shift perturbation; Glycosaminoglycans; HADDOCK; Hyaluronan

Link modules are domains found in proteins that are known primarily for their interaction with the non-sulfated glycosaminoglycan (GAG) hyaluronan (HA), a large polysaccharide composed of glucuronic acid (GlcA) and N-acetyl glucosamine (GlcNAc) ubiquitously present in the extracellular matrix of vertebrates^{1, 2}. In many cases HA-protein interactions contribute to stabilization of the matrix and thus tissue homeostasis^{3, 4}; examples include the massive complexes formed between HA and chondroitin sulfate proteoglycans that provide load bearing properties to cartilage, skin and brain via hydration/extension of the chondroitin sulfate chains^{1, 2}. However, synthesis of new matrices and matrix remodeling also occurs during certain physiological processes, such as ovulation, as well as in response to tissue injury, inflammation and disease⁴⁻⁶. In these cases, HA along with the sulfated GAGs, chondroitin sulfate (CS; a copolymer of GlcA and N-acetyl galactosamine (GalNAc)) and heparan sulfate (HS; a copolymer of GlcA or iduronic acid (IdoA) and GlcNAc), are known to play a key role in dictating tissue structure, cell fate and availability of extracellular signaling molecules in the extracellular matrix⁷⁻¹⁰. However, the structural basis for this expanded set of roles, particularly with respect to interactions of certain Link modules with GAGs other than HA, is less defined. Here we provide some of that basis using NMR methodology to study the interaction of the Link module, from human tumor necrosis factor-stimulated gene-6 (TSG-6), with a CS hexasaccharide.

TSG-6 is a multifunctional GAG-binding protein that has been implicated in matrix remodeling¹¹⁻¹³; it contains a single Link module that, unlike other members of this superfamily¹, binds to a wide range of sulfated GAGs (CS, dermatan sulfate, heparin and HS) in addition to HA¹⁴. The interaction of TSG-6 with HA leads to the direct crosslinking and structural reorganization of the polysaccharide via HA-induced dimerization of TSG-6, which promotes the interaction of HA with its cell surface receptor CD44. Heparin can also dimerize the isolated Link module domain from human TSG-6, where this may regulate the activity of plasmin and thus the turnover of matrix¹⁵; moreover, TSG-6 can mediate the crosslinking of structural proteins in the matrix¹⁶.

TSG-6 is not constitutively expressed in most adult tissues, but is secreted in response to inflammatory signals and during ovulation^{11, 12, 17-20}. In both of these contexts TSG-6 has

been found to catalyze the covalent transfer of heavy chains (HC) from the CS proteoglycan, inter- α -inhibitor (IaI), onto HA^{21, 22}; IaI has an unusual structure in that its two HCs are covalently linked via ester bonds to the CS chain of the bikunin core protein²³. The TSG-6 mediated formation of HA•HC complexes, which occurs via a covalent HC•TSG-6 intermediate²¹, is essential for murine ovulation^{24–26}; here attachment of HCs to HA facilitates the crosslinking of HA chains via their binding to pentraxin-3¹⁶, allowing the formation of a viscoelastic matrix around the oocyte just prior to its release into the oviduct. HC•HA complexes are also formed at sites of inflammation (e.g. in the synovial fluids from arthritis sufferers²⁷), where the HC transferase activity of TSG-6 has been found to be predictive of osteoarthritis severity and the need for joint replacement²⁸. The transfer of HCs onto HA likely involves the interaction of TSG-6 with the CS chain of IaI, i.e. during the formation of the HC•TSG-6 intermediate^{22, 29}.

The Link module of human TSG-6 (Link_TSG6) is well studied and has become a model for Link module interactions in other HA-binding proteins^{30–34}. Link_TSG6 is a small domain comprised of approximately 95 amino acids, and X-ray crystallography and solution NMR structures exist^{35, 33, 36}. Its interaction with hyaluronan (HA) has been extensively characterized by a combination of NMR and site-directed mutagenesis and likely involves residues K11, Y12, H45, V57, Y59, P60, I61, K63, F70, I76, Y78, R81 and W88^{30–36}. Heparin/HS binds to a site different from that of HA, namely one comprising residues K20, K34, K41 and K54^{15, 33}. Less information is known about where CS binds on TSG-6 Link module^{14, 37–39}. However, competition assays between CS and HA and binding assays with chimeric chondroitin and HA oligosaccharides (that are non-sulfated) suggest that CS might interact with the HA rather than the HS binding surface^{34, 37, 39}; it should be noted, however, that HA and heparin also compete for binding despite interacting at distinct sites, but this is believed to be mediated by an allosteric mechanism^{15, 33}. Therefore, it is unclear where CS binds and indeed whether a sulfated region of a CS chain could be accommodated in the Link module HA-binding groove without steric interference.

Sulfated GAGs are extremely heterogeneous materials; for CS and HS this is primarily due to the possibility of sulfation at a number of distinct positions on the constituent sugar rings^{7, 10}. HS can be N-sulfated, with the sulfate replacing the acetyl groups on the initial GlcNAc residues, as well as O-sulfated on the 3 and 6 positions of GlcNAc and the 2 position of IdoA residues. CS is less heterogeneous than HS with possibilities for sulfation at the 4 and 6 positions of the GalNAc residue in vertebrates.

A hexasaccharide of CS is an ideal way to study the interaction with Link_TSG6 since it is large enough to fill known binding sites and small enough to allow isolation or synthesis of a single well-defined oligosaccharide. However, even a hexasaccharide of CS has 64 possible sulfation patterns. Isolating a specific oligomer was overcome here by choosing a polymer that is rich in 4-O-sulfation (CS-A), digesting it with lyase or hydrolase enzymes, and purifying digests using size separation and ion exchange chromatography. This method allowed for the isolation of four CS candidates: C444S, C444S, C664S and C664S (in which the indicates the unsaturated lyase product and the numbers indicate the sites of sulfation on the three GalNAc residues, numbered from the non-reducing end). C444S was selected for structural studies of the Link module since it was available in larger quantities

and 4-sulfated CS binds with a higher affinity than the 6-sulfated hexamers (see Supplementary data), i.e. with a similar affinity to HA^{37, 38}.

NMR provides a reasonable approach to the structural characterization of protein-ligand complexes such as that of C444S with Link_TSG6. Measuring chemical shift perturbations in ¹H-¹⁵N HSQC spectra of ¹⁵N-labeled proteins provides one means of identifying amino acids that are potentially involved in binding. The structural information is very qualitative⁴⁰. However, once NMR resonance assignments are made, data are easily collected on relatively small amounts of sample, and analysis of shifts as a function of ligand concentration can provide binding constants. Here we supplement such chemical shift data with paramagnetic relaxation enhancement (PRE) data using a tagged version of the

C444S oligomer carrying a TEMPO group at the reducing end. The structural data is more quantitative and provides a definitive location for the reducing terminus of the hexasaccharide. Residual dipolar couplings (RDCs) between pairs of magnetic nuclei such as ¹H-¹⁵N pairs on amide groups of a protein backbone can also be used to provide quantitative data about changes in protein conformation through their dependence on the angles internuclear vectors make with the NMR magnetic field in partially ordered media^{41, 42}; they also provide information on the structure of dimers if these exist in solution. Here we used RDCs and chemical shift perturbation, along with computer modeling in order to produce a structure for the Link_TSG6: C444S complex. Our studies employed the docking program, HADDOCK, which incorporates a variety of quantitative and qualitative data in a simulated annealing search for a best structure for the complex^{43, 44}.

We will show below that the Link module from human TSG-6 has a primary CS interaction surface that utilizes the same binding groove as HA. We also show that ligand binding induces a conformational change in the Link module, that a dimer is formed (confirmed by analytical ultracentrifugation) and that a second CS-binding site partially overlaps the heparin-binding site.

METHODS

Expression, purification and refolding processes of Link_TSG6

Expression and refolding followed published procedures^{45, 46}. More specifically, the engineered gene of Link_TSG6 in the pRK172 expression vector was transformed into expression host BL21(DE3)pLysS. 10 ml starter cultures were grown overnight in LB medium containing 100µg/ml ampicillin. Grown cells were then transferred to 1 L M9 medium containing isotopes, ¹⁵NH₄Cl, ¹³C-glucose from CIL (Cambridge Isotope Laboratories) and protein expression was induced by adding IPTG to 0.1 mM when the OD_{600nm} reached 0.4. Cells were harvested 4 h after induction by centrifugation for 15 min at 5000g and stored at 20 °C in lysis buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl, and 1mM EDTA). Cells were lysed by three passages through a French press and cells and inclusion bodies were gathered by centrifugation for 45 min at 20000g.⁴⁵ Inclusion bodies were solubilized in 6M Guanidine-HCl containing 50mM Tris-HCl, pH 8.0, and 100 mM DTT and the sample was loaded onto a Superdex 75 (16 × 260 mm, Pharmacia) exclusion column, equilibrated and run with a flow rate at 1ml/min in the same buffer without DTT. Fractions containing the target protein were injected onto a C4 column (10 × 250 mm, YMC

America Inc), equilibrated with H₂O containing 0.01% TFA. After 5 min of washing with the same solvent at 3 ml/min, the protein was eluted over 40 min using a linear gradient from 0 to 80% acetonitrile containing 0.1% TFA. The eluent was monitored at 220 nm continuously. Eluent containing the denatured Link module was collected and lyophilized.

To accomplish the folding of Link_TSG6, the protein was re-suspended in 50 mM ammonium acetate at pH 6.0 at a concentration of 500 µg/ml (40 ml final volume). A 100-fold molar excess of β-mercaptoethanol was added, and the refolding solution was incubated at 25 °C for 2 days under aerobic condition without stirring. After 2 days the refolding solution was stored at the 4 °C for 5 additional days. The folded Link_TSG6 protein was purified by HPLC in the same manner as described above. The overall yield was approximately 10-20 mg/L.

Preparation of hexasaccharides of chondroitin sulfate

The method described below is based on the works described in Pomin et al.⁴⁷. 150 mg of CS-A from bovine trachea (btCS-A) was incubated with either 5 IU of chondroitin C lyase from *Flavobacterium heparinum* or 0.33 IU of chondroitin ABC lyase from *Proteus vulgaris* in 5 ml digestion buffer (50 mM Tris-HCl, pH 8.0, 150 mM sodium acetate, 100 µg/ml of BSA) at 37 °C for 36 h and 150 min, respectively. Hyaluronidase digestion of btCS-A was carried out by incubation of 150 mg of btCS-A with 10 mg of enzyme in 3 ml of digestion buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.0) at 37 °C for 48 h. The digested samples were heated to quench enzyme activities and subjected to separation on a Bio-Gel P-10, size exclusion column (15 × 1200 mm, Bio-Rad Life Science) with elution buffer (1 M NaCl, 10% ethanol) at a flow rate of 1.7 ml/15 min/fraction. Eluent was monitored at 232 nm and separated peaks in the chromatogram were desalted on a Sephadex G-15 column (10 × 500 mm, Sigma-Aldrich Co) and examined by mass spectrometry (MS) to identify those peaks having molecular weights corresponding to hexasaccharides. To eliminate complexities from anomeric equilibrium at the reducing end, most hexasaccharides were reduced in the presence of one equivalent of sodium borohydride (NaBH₄) in 1 ml of water for 3 h and the reaction was quenched by adding a molar equivalent of acetic acid for 1 h in an ice-bath, followed by desalting. Hexasaccharides were then purified on a strong anion exchange column, SAX (2.5 × 10 mm, 5 µm, Waters Corporation) using a linear gradient from 25 to 45% of 2 M NaCl, 25 mM phosphate, pH 4.5, for 50 min at a flow rate of 3 ml/min. Each peak from the SAX chromatography was desalted and lyophilized to yield 1 mg for C664S, 6.8 mg for C664S, 4.3 mg for C444S and 10.6 mg for C444S. Structures were identified by NMR spectroscopy and judged to be approximately 90% homogeneous.

Synthesis of C444S-TEMPO

4-amino-TEMPO (4-amino-2,2,6,6-tetramethylpiperidine-1-oxyl) was purchased from Acros Organics. The reductive amination reaction was carried out on a 400 µl sample containing 1 mg non-reduced C444S, 73 mM of 4-amino-TEMPO and 250 mM sodium cyanoborohydride in 80% MeOH at 65 °C for 3h. The sample was desalted and further purification was performed using SAX-HPLC monitored at 232nm. The yield was approximately 0.5 mg.

NMR spectroscopy

NMR spectra for titrations of protein with several isomers of CS were recorded at 25 °C on Varian 800- and 900-MHz spectrometers and processed and analyzed with NMRPipe and Sparky^{48, 49}. Titration experiments were carried out combining 0.2–0.5 mM uniformly ¹⁵N-labeled Link module with increasing amounts of CS oligomers (C664S, C444S, C664S, and C444S) in 50 mM MES buffer, pH 6.0, 0.02% NaN₃, 10% D₂O. A series of ¹H-¹⁵N HSQC spectra were recorded after adding sugar at 0.096, 0.192, 0.29, 0.38, and 0.48 mM to the protein sample. The degree of chemical shift change for amide proton and nitrogen resonances was calculated using the empirical formula, $\delta = [(\delta_{\text{HN}})^2 + (\delta_{\text{N}} \times 0.12)^2]^{1/2}$ where δ_{HN} and δ_{N} are the observed chemical shift changes for ¹H and ¹⁵N, respectively. The weighting factor of 0.12 reflects the difference in chemical shift dispersion of ¹H and ¹⁵N in folded proteins⁵⁰. For relaxation experiments using a TEMPO labeled version of C444S, a 0.17 mM ¹⁵N-labeled Link_TSG6 sample was prepared and spectra were recorded at 25 °C on a Varian 800-MHz spectrometer.

RDCs were measured on a 0.5 mM ¹³C and ¹⁵N labeled Link_TSG6 sample in the presence of 1 mM C444S aligned in 5 % (w/v) stretched neutral polyacrylamide gel. Gels were cast initially in a 4.5 mm diameter glass tube overnight for polymerization. The polymerized gels were washed for two or three cycles in deionized water overnight, followed by a washing with protein buffer to equilibrate the pH. Finally, the gels were washed again with deionized water to remove residual protein buffer. Swollen gels were then trimmed to a length of 26 mm and dried at room temperature for 2 days. The dried gel was placed in the upper stage of an NMR tube and protein sample was added to cover the gel and let it swell for 2 days. Spectra were recorded using a two-stage NMR tube as described in Liu et al.⁵¹. RDC values were measured using a ¹H-¹⁵N HSQC IPAP (in-phase and anti-phase) experiment under both isotropic and anisotropic conditions⁵².

Complete backbone and side chain resonance assignments were obtained with 0.5 mM ¹³C and ¹⁵N enriched Link_TSG6 in the presence of 0.6 mM C444S using the following heteronuclear 2D and 3D experiments: ¹H-¹⁵N HSQC, CBCA(CO)NH, HNCACB, HNCO and HCCH-TOCSY^{53, 54}. All spectra for assignment were collected at 25 °C on Varian 600- and 900-MHz spectrometers. ¹⁵N, and ¹³C-edited NOESY-HSQC were collected with 150 and 140 ms mixing time, respectively, on the same sample at 800-MHz and used for structure determination^{55, 56}.

In order to measure ¹⁵N T₁ and T₂ relaxation rates and evaluate the level of dimerization, spectra were recorded at 25 °C on a Varian 800-MHz spectrometer using standard sequences from BioPack (Varian/Agilent). ¹⁵N T₁ values were measured from the spectrum with different delay times, T = 10, 20, 30, 40, 50, 70, 90, 110, 150, 200, 250, 300, 400, and 500 ms. T₂ values were determined from spectra recorded with delays T = 10, 30, 50, 70, 90, 110, and 150 ms. T₁ and T₂ values were extracted from log plots of peak intensities in 2D spectra as a function of delay duration.

Analytical Ultracentrifugation (AUC)

All experiments were performed in 50 mM MES (pH 6.0) using an XL-A ultracentrifuge with an An60Ti -4 -hole rotor fitted with a six-sector epon-filled centerpiece with quartz glass windows. Equilibrium sedimentation was performed at 20°C, using rotor speeds of 22,000, 30,000 and 36,000 rpm scanning at 290 nm after equilibrium was reached (at 18 hours). Association kinetics were performed by using Link_TSG6 at 50 μ M and varying the concentration of chondroitin sulfate (C444S) between 50 and 1000 μ M. Analysis of the data was performed with non-linear regression using the Heteroanalysis program⁵⁷ using either an ideal non-associating model or monomer-dimer self-association model. For self-associating system analysis monomeric molecular weight of Link_TSG6 was fixed at 10.922 kDa and a molar extinction coefficient of 11976 (A_{290}).

NMR solution structure calculation

All NOESY crosspeaks were picked using Sparky and assigned manually followed by several rounds of automatic assignment. 1329 NOEs and 135 backbone torsion angle constraints derived from TALOS⁵⁸ using the assigned chemical shifts of HA, CA, CB, CO, and N were used to calculate an initial solution structure. CYANA version 3.0 calculations were performed and 20 structures with the lowest energies were selected for analysis and further refinement^{59, 60}. In the refinement stage, residual dipolar couplings were incorporated starting with Da and Rh values calculated from principle order parameters determined in REDCAT⁶¹ (−12.15 and 0.433, respectively). Structural refinement was then carried out using NOE distances, dihedral angles, and RDC values with Xplor-NIH⁶². Final ensemble contained 20 structures.

HADDOCK

HADDOCK version 2.1 was used for protein-protein and protein-ligand docking⁶³. The newly determined NMR structure of the protein was used for docking. The C444S structure was built in Glycam Biomolecular Builder⁶⁴ (also <http://dev.glycam.org>) and the TEMPO group was added manually using UCSF Chimera⁶⁵. Residues H4, R5, E6, A7, Y12, H45 and H96, were set as highly ambiguous interaction restraints (AIRs) for protein-protein docking. Residues, C47-A49, G58, K63-G65, and C68-I76, were used as AIRs for dimer- C444S-TEMPO docking. In addition, PRE data from spin labeled ligands were used as unambiguous interaction restraints for dimer- C444S-TEMPO docking. Initially 1,000 structures were determined by rigid-body docking. Then simulated annealing (SA) was carried out with the 200 lowest-energy structures using default force field parameters. Specified residues were allowed to be fully or semi flexible during the final stages of simulated annealing. For ligand docking C444S-TEMPO was allowed to be fully flexible and residue segments, of the protein 45–49, 63–65, and 68–76, were allowed to be semiflexible. For monomer-monomer docking to produce a dimer protein, residues 9–12 near the dimer interface were specified as semiflexible during the simulated annealing. Semiflexible residues were allowed to have side chains move during the SA. Fully flexible residues were allowed to be fully flexible throughout the entire docking protocol except for the rigid body minimization. RDC restraints were given during docking simulations and a tensor was included in the structures calculations. Finally the structures were subjected to a

water refinement stage. Models were clustered based on similarity of poses and ranked based on total score. The top five in each cluster were chosen for display and the lowest energy model of the top three scores was chosen for more detailed analysis. The best structures were chosen from the refined structures based on lowest total energy and docking score.

RESULTS

NMR-monitored titration of Link_TSG6 with CS

The isomers of the CS hexasaccharide (GlcA-GalNAc-GlcA-GalNAc-GlcA-GalNAc-ol; ol stands for reduced sugar with open ring at the reducing end) selected for our studies included, C444S, C444S, C664S and C664S. CS binding sites on Link_TSG6 were mapped using chemical shift perturbations in ^1H - ^{15}N HSQC spectra as a function of ligand concentration up to a twofold molar excess over protein concentration. Figure 1 shows the superimposition of the ^1H - ^{15}N HSQC spectra of 0.24 mM Link_TSG6 in the absence and presence of 0.096, 0.192, 0.29, 0.38, and 0.48 mM of the C444S hexasaccharide. Assignments of crosspeaks in the HSQC spectra were made using triple resonance experiments on a ^{15}N , ^{13}C labeled sample; assignments for the most significantly perturbed crosspeaks are indicated on Figure 1.

The chemical shift changes involved can be conveniently summarized by combining shifts in the two dimensions, ^1H and ^{15}N and plotting the maximum variations as a function of residue number (Figure 2). The residues with the largest shift changes are those most likely to be involved in ligand binding, and the extremes in chemical shift for these residues can be associated with shifts in uncomplexed and complexed states. An equivalent plot has been generated for $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ resonances (Figure S1 in Supplementary Data), and this shows large perturbations for a similar set of residues. Essentially identical behavior was observed in ^1H - ^{15}N HSQC spectra on titration with other hexasaccharides (Figures S2 and S3 in Supplementary Data).

Three distinct types of chemical shift change were observed; these correspond to exchange between complexed and uncomplexed states on slow, intermediate and fast timescales relative to the reciprocal of the chemical shift change in Hz. The color-coded bars in Figure 2A represent the residues involved in slow and fast exchange and the residue positions are indicated on the ribbon diagram of the Link_TSG6 structure in Figure 2B. In the first case, when ligand is added, crosspeaks progressively decrease in intensity as other peaks at nearby positions progressively increase in intensity. Crosspeaks belonging to residues H4, L14, H45, C47, A48, A49, G58, K63, C68, K72, G74, I76, and D89 fall into this category. A few crosspeaks belonging to G65, F70, and G71 also exhibited similar behavior (data not highlighted); in this case, the crosspeaks disappeared but could not be correlated with newly appearing peaks due to overlap or additional broadening in the complexed state.

Since intensity changes in the limit of slow exchange reflect the population of complexed and uncomplexed states, they can be used to determine a dissociation constant. Figure 3A shows a fit to the normalized average intensity changes for residues A48, G58, and G74. Equation 1 below describes the intensity change in terms of the normalized concentration of

the complex ($[C]/[PT]$). This depends on the binding constant, K_1 , the total protein concentration, $[PT]$, and the total ligand concentration, $[LT]$.

$$\frac{[C]}{[PT]} = \frac{1}{2} \left(\frac{1}{K_1[PT]} + \frac{[LT]}{[PT]} + 1 \right) \pm \frac{1}{2[PT]} \sqrt{\left(\frac{1}{K_1} + [LT] + [PT] \right)^2 - 4[PT][LT]} \quad [\text{Equation 1}]$$

$[PT] - [C]/[PT]$ gives the corresponding curve for free protein, $[P]/[PT]$, which is also shown in Figure 3A. In the limit of slow exchange the amplitude of a particular cross peak is proportional to $[P]$ or $[C]$. However, lifetime broadening also occurs as exchange rates (KE) increase, leading to an additional reduction in amplitude at low concentrations of $[C]$ or $[P]$ in proportion to KE, the effective exchange rate at the titration midpoint in s^{-1} . This effect is easily added to the simulation by scaling amplitudes by the ratio of the initial line width (20 Hz) to the linewidth in the presence of lifetime broadening, more specifically the initial line width plus KE/π times the fraction in complexed or uncomplexed states. We have also compensated for the linewidth increase due to dimerization (see below) by scaling up amplitudes of experimental complex peaks by a factor of 1.4, which is based on the observed linewidth for those peaks.

Simulations of $[C]/[PT]$ and $[P]/[PT]$ as a function of $[L]$ were done for a grid of K_1 and KE parameters, and the best fit simulations are plotted in Figure 3A. While amplitude changes depart somewhat from ideal behavior due to factors such as intensity losses during HSQC transfers due to broadened peaks, which are not included in the simulation, it was possible to find a reasonable fit with a binding constant, K_1 of $7.5 \times 10^4 \text{ M}^{-1}$, or equivalently, a dissociation constant, K_d , of $13 \text{ }\mu\text{M}$, and a KE of 30 s^{-1} for the C444S-Link_TSG6 interaction. The fittings to the normalized average intensity changes for C664S and C664S are provided in Figure S4 of Supplementary Data.

For a second class of perturbed peaks, crosspeaks displayed some chemical shift change and signal broadening as the ligand concentration increased. In most cases, this resulted in the disappearance of crosspeaks when the total ligand concentration was approximately one-half of the total protein concentration. These crosspeaks then reappeared at much higher ligand concentration. This is characteristic of an exchange process occurring on an intermediate timescale with respect to the reciprocal of chemical shift changes. Residues E18, G50, N67, I75 and W88 exhibited this behavior. The changes in chemical shifts were, in general, smaller than those for residues showing slow exchange, making them more susceptible to exchange broadening. The residues are in the same structural area where slow exchanging peaks are found. Hence, they are likely involved in the same ligand binding process.

A third type of behavior is characterized by peaks exhibiting small changes in chemical shift and line width at lower ligand concentrations, followed by larger continuous chemical shift changes at higher ligand concentrations. Crosspeaks for K11, E24, H29, L30, Y33, W51, A53, Y91 and Y93 belong in this category. Their behavior is a special case of fast exchange with the onset of this second weaker binding event occurring only after binding at the high affinity site is completed. Under rapid exchange, chemical shifts are population-weighted averages of complexed and uncomplexed states. The set of equations 2–4 below describe the

change in concentrations of various species as a function of ligand concentration. They are based on the assumption that only the products of slow exchange are capable of participating in the fast exchange process and the fast exchange process involves binding to a dimer of the initial complex (see results section to follow).

$$[C2] = K_2 \times [L] \times K_3 \times (K_1 \times [L] \times [P])^2 \quad \text{[Equation 2]}$$

$$[PT] = [P] + 2 \times [C2] + K_1 \times [L] \times [P] \quad \text{[Equation 3]}$$

$$[LT] = [L] + K_1 \times [L] \times [P] + 3 \times [C2] \quad \text{[Equation 4]}$$

Here [C2] is the concentration of the second complex. [PT] and [LT] are total concentration of protein and ligand, respectively. K_1 represents a binding constant for slow exchange process, K_3 is a protein dimerization constant, and K_2 represents a binding constant for the second binding event. K_1 was set to $7.5 \times 10^4 \text{ M}^{-1}$ based on the fit to slow exchange data, K_3 was set to 10^4 M^{-1} based on an AUC experiment (see below) and [PT] was set to 0.24 mM. The chemical shift change is proportional to the fraction of protein in the second complex ($[C2]/[PT]$). The system of equations was solved using the Maple program, and simulations of the scaled chemical shift change were carried out with a series of binding constants, K_2 , in order to find the best fit to the experimental data. The resulting best curve is shown along with experimental data in Figure 3B. The binding constant extracted is $2 \times 10^4 \text{ M}^{-1}$ (i.e. a K_d of 50 μM).

Binding of Link_TSG6 with a TEMPO derivative of C444S

In order to better define the position of the bound C444S, a paramagnetic version of C444S (C444S-TEMPO), which includes a nitroxide carrying TEMPO group attached to the reducing end, was used to generate additional data (sugar sequence; 1GlcA-2GalNAc4S-3GlcA-4GalNAc4S-5GlcA-6GalNAc4S1N-TEMPO). Figure 4 shows a comparison of a sample of 0.17 mM Link_TSG6 with 0.34 mM C444S-TEMPO in oxidized (4A, paramagnetic) and reduced (4B, diamagnetic) states; the reduced state was achieved by the addition of twelve equivalents of ascorbic acid⁶⁶. With few exceptions, chemical shift perturbations were identical for C444S with and without the attached TEMPO group suggesting that similar modes of binding exist. The time scales of exchange also appear to be similar. For example, A48 still shows slow exchange when it interacts with C444S-TEMPO.

In the oxidized state certain crosspeaks lose intensity due to the paramagnetically enhanced decay of transverse magnetization during the transfer and refocusing periods of the HSQC experiment. The changes in intensity between reduced and oxidized TEMPO derivatives can

be converted to distances between the nitroxide oxygen and a residue of interest using equation 5 below^{67–69}.

$$\ln \left(\frac{I_{red}}{I_{ox}} \right) = f \times t \times \left(\frac{K}{r^6} \right) \times \left(4\tau_c + \frac{3\tau_c}{1 + \omega_H^2 \tau_c^2} \right) \quad [\text{Equation 5}]$$

Here, I_{red} and I_{ox} are the peak intensities of the crosspeaks in the presence of reduced and oxidized TEMPO derivatives, f is the fraction of the protein bound to ligand, and t is the total time during INEPT and the refocusing periods of the ^1H - ^{15}N HSQC pulse sequence ($t=9.78$ ms). K is a constant related to spin properties of the system ($K=1/15 \times S(S+1)\gamma^2 g^2 \beta^2 = 1.23 \times 10^{-23} \text{ cm}^6 \text{S}^{-2}$), r is the distance between the nitroxide and amide proton of the crosspeak of interest, τ_c is the correlation time for tumbling of the protein-ligand complex, and ω_H ($800 \text{ MHz} \times 2\pi$) is the precession frequency for the amide proton. Given all values, and taking f and τ_c to be 1 and 10.15 ns, respectively, the distances shown in Table 1 were obtained.

The distances relate only to the positioning of the terminal TEMPO group. The amides of residues G79, I80, and N83 are most affected by the presence of the TEMPO group, with the crosspeak for I80 completely disappearing. Even though a direct correlation with residues showing chemical shift perturbation was not necessarily expected, G79 and I80 are near the 74–76 segment showing slow exchange, suggesting that distances involving these residues would be useful for positioning a ligand in the slow exchange binding site. Given that the TEMPO is attached to the reducing end of the hexasaccharide, data are consistent with the findings by Blundell et al., 2003 and Higman et al., 2014, showing the reducing end of bound HA to be close to Y78 and R81^{34, 36}.

Link_TSG6 forms a dimer in the presence of C444S

It is possible that some of the observed chemical shift perturbations arise from interactions that do not directly involve the ligand, but are interactions in remote parts of the protein that are induced by ligand binding. One possibility is ligand induced dimer formation with perturbed residues being involved in a dimer interface. NMR spin relaxation measurements provide one way of assessing the possibility of dimer formation. R_1 and R_2 relaxation rates for ^{15}N nuclei in amide groups were measured in free and C444S-bound states. The ratio of R_2/R_1 yields a rotational correlation time for a macromolecule, provided the amide sites measured do not experience substantial internal motion. R_1 , R_2 , and R_2/R_1 ratios are reported in Figure 5. Basic equations for this ratio⁷⁰ can be reduced to an approximate relation that can be solved for a correlation time, $R_2/R_1 - 7/6 \approx 2/3 (\tau_c \times \omega_N)^2$ where ω_N is the precession frequency for ^{15}N . Analysis yields correlation times of 6 and 11 ns for uncomplexed and complexed proteins respectively. At 25 °C in dilute aqueous solution the correlation times in ns derived from sites in rigid parts of the protein are expected to approximately equal half the molecular weight in kDa (derived from the Stokes formula for rotational diffusion). Hence, the value of 6 ns for Link_TSG6 at a concentration of 0.26 mM in the absence of ligand is in reasonable agreement with that expected for a monomer of

10.9 kDa. In the presence of C444S, the rise to 11 ns, is much greater than that expected for the addition of ligand to the protein and is more consistent with dimer formation.

Analytical ultracentrifugation analysis of Link_TSG6 in the absence and presence of

C444S provided further evidence for CS-induced dimerization. Link_TSG6 without any added C444S gave a sedimentation equilibrium profile that fit well to an ideal single species model showing that in the absence of CS hexasaccharide the protein is monomeric; all experiments were conducted at three different rotor speeds; data at 30,000 rpm are illustrated in Table 2 and Supplementary Figure S5. In the presence of C444S data were fit to a monomer-dimer self association model. The molecular mass of the protein species increased dose-dependently with increasing concentrations of the added hexasaccharide (at 1:1, 1:10 and 1:20 molar ratios; Table 2). Thus, this AUC analysis is consistent with dimerization of Link_TSG6 in the presence of C444S; an apparent dissociation constant of 151 μ M was derived by non-linear regression of the data at the 1:20 molar ratio.

Since C444S addition induces dimer formation, it is likely that some of the ligand induced chemical shifts attributed to C444S-protein interactions are the direct result of dimer formation. We therefore attempted to gain information about possible dimer contacts by examining ^1H - ^{15}N HSQC spectra in the absence of C444S, but at higher concentrations. Comparing chemical shifts at 0.1 mM and 0.5 mM there were a number of small but measureable changes, including H4, R5, E6, A7, Y12, H29, L30, H45, V46, Y91 and H96 (Figure S6 in Supplementary data); two of these, H29 and L30, were discounted since these were apart from other perturbed residues and the fact that histidine ionization can be altered with slight changes in pH and ionic strength as the samples are concentrated^{71, 72}. The remaining perturbed resonances were deemed to be useful for the determination of a dimer structure; however, some of these perturbations do correspond to those seen with the addition of C444S and therefore it will be necessary to eliminate the use of these in the determination of the ligand binding site.

Link_TSG6: C444S solution NMR structure

The existence of a dimer and the possibility that ligand binding induces additional structural changes suggest that refinement of a structure in the presence of C444S would be worthwhile. We, therefore, pursued solution structure determinations using conventional triple resonance assignment methods combined with NOE, dihedral angle, and RDC restraints. A statistical summary of the data is included in Table 3. The initial structure for a monomer of Link_TSG6 in the presence of C444S was generated using CYANA version 3 with 1198 manually assigned NOEs out of 1329 total NOEs and 135 dihedral angles predicted from TALOS⁷³. In addition, two disulfide bonds were created and restrained to normal bond distances, based on previous disulfide bond mapping and their existence in the crystal structures^{33, 45}. The structure was then further refined using XPLOR-NIH with additional NOEs iteratively assigned based on initial structures from CYANA calculations, and with added RDC restraints. The final stage included refinement in the presence of discrete water molecules as solvent. 20 models having the lowest total energy were chosen and are represented in Figure 6. This model has two identifiable α -helices and two small β -sheet-like structures. Using default secondary structure parameters the terminal and central

paramagnetic effects and slow exchange chemical shift effects, data from both effects were combined to generate a model that provides insight into the higher affinity C444S binding site. The following ambiguous interaction restraints were used: C47-A49, G58, K63-G65, and C68-I76 along with all distance restraints determined with C444S-TEMPO. Distance restraints coming from PREs were used with lower limits set to 3 Å and upper limits set to the measured distance plus 3 Å. 46 RDC restraints on protein ^{15}N - ^1H vectors were provided to allow additional adjustments of loop regions designated as flexible during final stages of docking (residues 45–49, 63–65, and 68–76). The top 10 scoring complexes from HADDOCK were selected and clustered. These are shown in Figures 7A and 7B. One cluster (Figure 7A) shows C444S-TEMPO sitting predominantly at a single site on the protein surface involving residues 62–76 but with a broad range of conformations. The other cluster (Figure 7B) shows the C444S-TEMPO docked into a shallow groove overlapping the HA-binding site (see below).

The average energies for the two clusters are similar, i.e. –6986.6 and –6940.4 kcal/Mol, respectively, but the second cluster makes better contacts with the perturbed resonances (H4, L14, H45, C47, A48, A49, G58, K63, C68, K72, G74, I76, and D89). Also, the model with the lowest molecular energy (sum of E_{tot} , E_{bond} , E_{angle} , E_{improper} , E_{dihedral} , E_{vdw} , and E_{elec}) among the top 3 scoring models, is a member of cluster 2; it also shows the lowest electrostatic energy (E_{elec}). Therefore, this model was chosen for further analysis/refinement of specific interactions between the protein and C444S because it provides a reasonable representation of the oligosaccharide sitting within its primary binding site (Figure 8). 6 potential hydrogen bonds ($\text{H}\cdots\text{O}$ distance < 2.5 Å) were observed between the protein and C444S-TEMPO (1GlcA-O5:H45-HE2, 2GalNAc4S-O3:H45-HE2, 3GlcA-O5:N67-HD21, 3GlcA-OE2:N67-HD21, 3GlcA-OE1H:C68SG and 5GlcA-O5:G69-NH). In addition, 115 atom pairs are within 0.2 Å of van der Waals contact distance across the protein-sugar interface. Residues K11, H45-A49, Y59, P64, N67-G59, F70 K72, I76 and Y78 were involved in these contacts. Sulfate groups on two GalNAc residues were in close contact with the side chain of positively charged amino acids (K11 and H45). The other sulfate group, which is in the opened ring produced on TEMPO attachment was located outside of the binding groove where it could make effective contact with water. The nitroxide on the TEMPO group showed distances to G79 (11.76 Å), I80 (12.58 Å), and N83 (9.39 Å) that are consistent with distance restraints from the PRE data.

The bound conformation of C444S differs from that expected in solution. Sugar rings all retain the expected $^4\text{C}_1$ chair conformation, however glycosidic torsion angles vary. The glycan forcefield parameters used in our docking procedure were not specifically optimized for glycans and this can contribute to the departures. Recently the use of an auxiliary energy function specific for glycosidic torsion angles has been proposed as a means to improve scoring of poses produced by docking programs that lack adequate treatment of glycan conformational energies⁷⁵. Evaluation of this function for the central three glycosidic bonds in the five poses in our second cluster show the selected conformer described above to have the lowest score (proper energy functions do not exist for the terminal residues) with the total penalty being less than 4 kcal. Also, departures from perturbations in glycosidic bond angles from solution conformations have been noted in HA-protein complexes^{34, 76}.

Interestingly, the position of the pronounced kink in the structure shown in Figure 8 is similar to that observed previously. The information obtained on the weaker, fast exchange site was not adequate to generate a detailed structural model for C444S binding. However, the perturbed sites are shown on the dimer structure in Figure 8A. The positioning is clearly separate from the primary site, allowing the possibility of simultaneous interactions at the two sites.

DISCUSSION

The investigations of the interactions of chondroitin sulfate hexasaccharides allow us to make several general observations: Chondroitin sulfate oligomers such as C444S bind to Link_TSG6 with moderate affinity and a preference for high 4-sulfate content. There are at least two different binding sites, one characterized by slow exchange and tight binding ($K_d = 13\mu\text{M}$) one by fast exchange and weaker binding ($K_d = 50\mu\text{M}$). And, the initial binding event promotes Link_TSG6 dimer formation. The studies have allowed the generation of a structural model for the dimer and the bound conformation of the C444S hexasaccharide. Here we discuss this model in light of previous work on binding of HA and heparin/HS to Link_TSG6, and the biological implications of our model.

The similarity of the slow exchange (higher affinity) binding site to that previously observed for HA oligomers is striking³⁴. The similarity can even be seen at the level of residues experiencing perturbed chemical shifts on addition of C444S. The following residues were previously identified as being in proximity to HA (i.e. lining the binding groove): K11, Y12, H45, V57, Y59, P60, I61, K63, F70, I76, Y78, R81 and W88^{30–36}; CH- π stacking interactions of saccharide rings against aromatic planes for Y59 and Y78 were suggested to play an important role^{32, 34, 36}. The list of residues with slow and intermediate exchange rates encompasses five of the 13 residues listed above. If we count residues that are one or two removed in sequence, the list expands to 8 of the 13 residues. Another piece of evidence to show C444S shares the HA binding site is the perturbation of C47 and C68 resonances. These residues, which form a disulfide bridge underlying the HA binding surface, have previously been suggested to change geometry on binding with a concomitant rearrangement of the $\beta 4$ - $\beta 5$ loop (residue 61–74) leading to this perturbation³⁶. If we add these to the list, 10 of 15 are encompassed. All of the perturbed residues for C444S are highlighted in Figure 9B and compared to those residues believed to be involved in the HA binding in Figure 9A. Even though C444S and HA share a binding groove, the residues perturbed by C444S do differ in some cases. Positively charged amino acids, such as H4, and K72 (that do not interact with HA^{31, 72}), in addition to H45 and K63 (that do³⁴), are perturbed indicating that they may be associating with sulfate groups on the GalNAcs of C444S.

The binding region of the protein also seems to experience some distinct structural adjustments on interaction with C444S when compared to HA. The overall differences are modest; a comparison of Link_TSG6 in its C444S- and HA-bound conformations yields an RMSD for backbone atoms of 1.34 Å. However, the large loop (residue 61–74) involved in binding C444S and HA shows a larger difference. As illustrated in Figure 10, the binding groove appears to be closed more tightly when C444S is present than when HA is present. Given the similar shift pattern seen with the C664S oligomers (Supplementary Figures S2

and S3), it is evident that CS hexasaccharides containing 6-sulfates can also be accommodated within this binding groove, albeit with lower affinity. The lower affinity is consistent with previous observations indicating that chondroitin-6-sulfate (C6S; also termed CS-C), unlike chondroitin-4-sulfate (C4S; also termed CS-A), is a poor competitor of HA^{37, 39}. Based on the model for Link_TSG6: C444S (Figure 8) it would appear that the presence of sulfates on the 6-position of GalNAc are unlikely to cause steric clashes with the protein. However, it is possible that 6-sulfates are not be able to make as favorable ionic interactions with positively charged residues in the protein, or perhaps the inherent differences in C4S and C6S chain flexibility^{77, 78} may disfavor the binding of C6S.

The second mode of interaction of C444S with the TSG-6 Link module, that exhibited fast exchange and a higher K_d of approximately 50 μ M, was more difficult to characterize. However, the residues perturbed in this interaction are near some of the residues previously identified as important for heparin binding (K20, K34, K41, and K54)^{15, 33}. As shown in Figure 2, the residues perturbed in this weaker interaction are physically separated from those exhibiting slow exchange. However, it is possible that longer CS chains could bind simultaneously to the two binding sites, which would increase the affinity of the interaction for longer polymers. This would also be expected have a large effect on the conformation of the polysaccharide by wrapping it around two sides of the Link module (see Figure 8A). This would cause a condensation of a CS chain (as we have seen for HA¹³) and thereby reduce the overall domain size of a CS proteoglycan. As noted previously^{13, 34}, such structural perturbation could drive matrix reorganization at inflammatory sites, where the TSG-6 protein is expressed. For example, TSG-6-mediated collapse of the CS chains of aggrecan could enhance the movement of this proteoglycan through the cartilage extracellular matrix aiding tissue maintenance and repair³⁴.

An additional discovery is that binding of C444S to the Link module induces dimer formation. The dimer structure has an interface distinctly different from the possible dimer contacts suggested by the crystal structure determined in the absence of ligands³³. However, the reduction in solvent accessible surface areas for the three possible dimers in the non-liganded crystal structure are not significantly different (625, 884, 612 $\text{\AA}^2$) from that observed here 723 $\text{\AA}^2$. Another aspect of the dimer structure worth noting is the positioning of the C-termini that must connect to a CUB domain in intact TSG-6. These extend upward from the cleft between monomers seen in Figure 7. A crystal structure of the CUB domain has recently been published⁷⁹, and without altering the structure of either domain we have been able to dock a pair of CUB modules to these termini with a minimal clash of surface residues (Supplementary Figure S7). The CUB modules are in close proximity to one another allowing for possible additional interactions in the context of a TSG-6 dimer. Based on our modeling the N-terminal peptide of TSG-6 could also potentially be accommodated within the CS-induced dimer, however, given that there are no structural data available for this region, further work would be needed to assess this possibility.

In our model, primary C444S binding sites exist on each monomer, nearly on opposite sides of the protein from the dimerization interface. Thus, the hexasaccharide ligands used here cannot play a bridging role in dimer formation. They could however lead to a reduction in repulsive electrostatic effects. Link_TSG6 is a positively charged domain ($pI = 9.48$;

calculated using the compute PI tool in ExPASy). The binding of a negatively charged ligand would then reduce the repulsion that exists in the unliganded monomers and allow dimerization. One consequence of dimerization is that multiple binding sites for CS will be brought together (with two on each monomer). Thus, it is possible that a polymeric CS chain might bind to the Link module dimer in a number of different ways, utilizing several different combinations of the four potential interaction surfaces. Importantly, some of these combinations would be anticipated to lead to CS crosslinking via the binding of two CS molecules to the Link module dimer; for example, each CS chain could associate with the high and low affinity sites as suggested above. Crosslinking of CS by TSG-6 could have a pronounced effect on the structural and functional properties of tissues such as brain, cartilage and skin that have a high CS content. For example, rigidification of matrix could alter cell phenotype through changes in mechanical sensing^{80–82}. The structural reorganization of CS could also modulate the binding of matrix-associated signaling molecules, such as the chemokine CCL5⁸³; interestingly, TSG-6 binds to CCL5⁸⁴ as well as CS, and therefore may be able to interfere with its presentation (e.g. on endothelial cells).

CS-induced dimerization of the TSG-6 Link module could also be of relevance to the covalent transfer of the heavy polypeptide chains of IaI onto HA; this is catalyzed by TSG-6²¹ and occurs via two sequential transesterification reactions, where the first of these transfers a heavy chain from CS onto TSG-6 to form a HC•TSG-6 intermediate. Importantly, there is evidence that the binding of TSG-6 to CS is involved in HC•TSG-6 formation^{22, 29, 85}. Thus, HC transfer might potentially involve dimerization of TSG-6 via its binding to CS facilitating a subsequent transfer of HC to HA bound to a second TSG-6 molecule. While this cannot be ruled out, it seems unlikely given the weak affinity for dimerization determined here by AUC ($K_d = 150 \mu\text{M}$), whereas HC•TSG-6 complexes can form readily when TSG-6 and IaI are present at low micromolar concentrations^{21, 22}. In any case, the finding that TSG-6 interacts with CS through its 'HA binding groove' will be helpful in further refining our molecular understanding of this important biological process (e.g. in the context of inflammation and ovulation).

There is also a precedent for the dimerization of Link_TSG6/TSG-6 on binding to GAGs. Heparin, for example, is known to induce dimer formation of Link_TSG6 in a manner that may potentiate the anti-plasmin activity of IaI¹⁵. In this case, it has been suggested that heparin may act as a bridge between the two Link modules by binding to continuous surface on the dimer, i.e. based on crystal contacts in the Link_TSG6 X-ray structure³³. This is clearly a different mechanism from that identified here for CS. More similar is the binding of HA to TSG-6, which induces dimerization of the full-length protein, where this leads to crosslinking and a dramatic collapse and stiffening of the HA network^{13, 16}. However, Link_TSG6 is not dimerized by HA, nor is its interaction with HA cooperative (as is the case for full-length TSG-6); it is the CUB module that likely mediates the protein-protein interaction between TSG-6 molecules¹³. Thus, the mechanisms underlying GAG-induced dimerization of TSG-6, and its Link module domain, appear to be distinct for CS, HA and heparin, indicating that TSG-6 may be able to play diverse roles in tissue remodeling. However, specific information related to how TSG-6 interacts with polymeric CS and whether this does indeed lead to crosslinking within extracellular matrix requires additional studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Vitor H. Pomin for his assistance in the preparation of the CS hexasaccharides, Laura Morris for assistance with the modeling and Dr. Charles D. Blundell for helpful advice on the manuscript.

Funding

This work was supported by a grant from the NIH Institute of General Medical Sciences P41GM103390 (JHP); it also benefited from instrumentation provided in part by grant S10RR027097. AJD acknowledges Arthritis Research UK for their funding (16539, 18472).

ABBREVIATIONS

NMR	nuclear magnetic resonance
AUC	analytical ultracentrifugation
TSG-6	tumor necrosis factor-stimulated gene-6
Link_TSG6	Link module from TSG-6
GAG	glycosaminoglycan
ECM	extracellular matrix
HA	hyaluronic acid
CS	chondroitin sulfate
C444S	lyase produced CS hexasaccharide sulfated at the 4-oxygen of each N-acetylgalactosamine
RDC	residual dipolar coupling
NOE	nuclear Overhauser effect
PRE	paramagnetic relaxation enhancement
HS	heparan sulfate
HC	heavy chain
HSQC	heteronuclear single quantum coherence
TEMPO	2,2,5,5-tetramethyl-1-piperidinyloxy
PDB	protein data bank
BMRB	biological magnetic resonance data bank

References

1. Day AJ, Prestwich GD. Hyaluronan-binding proteins: tying up the giant. *J Biol Chem.* 2002; 277:4585–4588. [PubMed: 11717315]
2. Day AJ, de la Motte CA. Hyaluronan cross-linking: a protective mechanism in inflammation? *Trends Immunol.* 2005; 26:637–643. [PubMed: 16214414]
3. Tammi MI, Day AJ, Turley EA. Hyaluronan and homeostasis: A balancing act. *Journal of Biological Chemistry.* 2002; 277:4581–4584. [PubMed: 11717316]
4. Monslow J, Govindaraju P, Pure E. Hyaluronan - a functional and structural sweet spot in the tissue microenvironment. *Frontiers in Immunology.* 2015; 6:231. [PubMed: 26029216]
5. Petrey AC, de la Motte CA. Hyaluronan, a crucial regulator of inflammation. *Frontiers in Immunology.* 2014; 5(101):1–13. [PubMed: 24474949]
6. Wells JM, Gagger A, Blalock JE. MMP generated matrikines. *Matrix Biology.* 2015; 44–46:122–129.
7. Kwok JCF, Warren P, Fawcett JW. Chondroitin sulfate: A key molecule in the brain matrix. *International Journal of Biochemistry & Cell Biology.* 2012; 44:582–586. [PubMed: 22265655]
8. Ayerst BI, Day AJ, Nurcombe V, Cool SM, Merry CLR. New strategies for cartilage regeneration exploiting selected glycosaminoglycans to enhance cell fate determination. *Biochemical Society transactions.* 2014; 42:703–709. [PubMed: 24849240]
9. Langford-Smith A, Day AJ, Bishop PN, Clark SJ. Complementing the sugar code: role of GAGs and sialic acid in complement regulation. *Frontiers in Immunology.* 2015; 6(25):1–7. [PubMed: 25657648]
10. Xu, D.; Esko, JD. Demystifying Heparan Sulfate-Protein Interactions. In: Kornberg, RD., editor. *Annual Review of Biochemistry.* Vol. 83. 2014. p. 129-157.
11. Milner CM, Day AJ. TSG-6: a multifunctional protein associated with inflammation. *J Cell Sci.* 2003; 116:1863–1873. [PubMed: 12692188]
12. Milner CM, Higman VA, Day AJ. TSG-6: a pluripotent inflammatory mediator? *Biochemical Society transactions.* 2006; 34:446–450. [PubMed: 16709183]
13. Baranova NS, Nileback E, Haller FM, Briggs DC, Svedhem S, Day AJ, Richter RP. The inflammation-associated protein TSG-6 cross-links hyaluronan via hyaluronan-induced TSG-6 oligomers. *J Biol Chem.* 2011; 286:25675–25686. [PubMed: 21596748]
14. Marson A, Robinson DE, Brookes PN, Mulloy B, Wiles M, Clark SJ, Fielder HL, Collinson LJ, Cain SA, Kielty CM, McArthur S, Buttle DJ, Short RD, Whittle JD, Day AJ. Development of a microtiter plate-based glycosaminoglycan array for the investigation of glycosaminoglycan-protein interactions. *Glycobiology.* 2009; 19:1537–1546. [PubMed: 19729381]
15. Mahoney DJ, Mulloy B, Forster MJ, Blundell CD, Fries E, Milner CM, Day AJ. Characterization of the interaction between tumor necrosis factor-stimulated gene-6 and heparin: implications for the inhibition of plasmin in extracellular matrix microenvironments. *J Biol Chem.* 2005; 280:27044–27055. [PubMed: 15917224]
16. Baranova NS, Inforzato A, Briggs DC, Tilakaratna V, Enghild JJ, Thakar D, Milner CM, Day AJ, Richter RP. Incorporation of pentraxin 3 into hyaluronan matrices is tightly regulated and promotes matrix cross-linking. *J Biol Chem.* 2014; 289:30481–30498. [PubMed: 25190808]
17. Wisniewski HG, Hua JC, Poppers DM, Naime D, Vilcek J, Cronstein BN. TNF/IL-1-inducible protein TSG-6 potentiates plasmin inhibition by inter-alpha-inhibitor and exerts a strong anti-inflammatory effect in vivo. *J Immunol.* 1996; 156:1609–1615. [PubMed: 8568267]
18. Fulop C, Kamath RV, Li Y, Otto JM, Salustri A, Olsen BR, Glant TT, Hascall VC. Coding sequence, exon-intron structure and chromosomal localization of murine TNF-stimulated gene 6 that is specifically expressed by expanding cumulus cell-oocyte complexes. *Gene.* 1997; 202:95–102. [PubMed: 9427551]
19. Carrette O, Nemade RV, Day AJ, Brickner A, Larsen WJ. TSG-6 is concentrated in the extracellular matrix of mouse cumulus oocyte complexes through hyaluronan and inter-alpha-inhibitor binding. *Biology of reproduction.* 2001; 65:301–308. [PubMed: 11420253]
20. Mukhopadhyay D, Hascall VC, Day AJ, Salustri A, Fulop C. Two distinct populations of tumor necrosis factor-stimulated gene-6 protein in the extracellular matrix of expanded mouse cumulus

- cell-oocyte complexes. Archives of biochemistry and biophysics. 2001; 394:173–181. [PubMed: 11594731]
21. Rugg MS, Willis AC, Mukhopadhyay D, Hascall VC, Fries E, Fulop C, Milner CM, Day AJ. Characterization of complexes formed between TSG-6 and inter-alpha-inhibitor that act as intermediates in the covalent transfer of heavy chains onto hyaluronan. J Biol Chem. 2005; 280:25674–25686. [PubMed: 15840581]
 22. Briggs DC, Birchenough HL, Ali T, Rugg M, Waltho JP, Ievoli E, Jowitt TA, Enghild JJ, Richter RP, Salustri A, Milner CM, Day AJ. Metal Ion-Dependent Heavy Chain Transfer Activity of TSG-6 Mediates Assembly of the Cumulus-Oocyte Matrix. J Biol Chem. 2015; 290:28708–28723. [PubMed: 26468290]
 23. Enghild JJ, Thogersen IB, Cheng F, Fransson LA, Roepstorff P, Rahbek-Nielsen H. Organization of the inter-alpha-inhibitor heavy chains on the chondroitin sulfate originating from Ser(10) of bikunin: posttranslational modification of IalphaI-derived bikunin. Biochemistry. 1999; 38:11804–11813. [PubMed: 10512637]
 24. Ochsner SA, Day AJ, Rugg MS, Breyer RM, Gomer RH, Richards JS. Disrupted function of tumor necrosis factor-alpha-stimulated gene 6 blocks cumulus cell-oocyte complex expansion. Endocrinology. 2003; 144:4376–4384. [PubMed: 12959984]
 25. Ochsner SA, Russell DL, Day AJ, Breyer RM, Richards JS. Decreased expression of tumor necrosis factor-alpha-stimulated gene 6 in cumulus cells of the cyclooxygenase-2 and EP2 null mice. Endocrinology. 2003; 144:1008–1019. [PubMed: 12586778]
 26. Fulop C, Szanto S, Mukhopadhyay D, Bardos T, Kamath RV, Rugg MS, Day AJ, Salustri A, Hascall VC, Glant TT, Mikecz K. Impaired cumulus mucification and female sterility in tumor necrosis factor-induced protein-6 deficient mice. Development. 2003; 130:2253–2261. [PubMed: 12668637]
 27. Yingsung W, Zhuo LS, Morgelin M, Yoneda M, Kida D, Watanabe H, Ishiguro N, Iwata H, Kimata K. Molecular heterogeneity of the SHAP-hyaluronan complex - Isolation and characterization of the complex in synovial fluid from patients with rheumatoid arthritis. Journal of Biological Chemistry. 2003; 278:32710–32718. [PubMed: 12799384]
 28. Wisniewski HG, Colon E, Liublińska V, Karia RJ, Stabler TV, Attur M, Abramson SB, Band PA, Kraus VB. TSG-6 activity as a novel biomarker of progression in knee osteoarthritis. Osteoarthritis and Cartilage. 2014; 22:235–241. [PubMed: 24333293]
 29. Lord MS, Day AJ, Youssef P, Zhuo L, Watanabe H, Caterson B, Whitelock JM. Sulfation of the Bikunin Chondroitin Sulfate Chain Determines Heavy Chain center dot Hyaluronan Complex Formation. Journal of Biological Chemistry. 2013; 288:22930–22941. [PubMed: 23801333]
 30. Kahmann JD, O'Brien R, Werner JM, Heinegard D, Ladbury JE, Campbell ID, Day AJ. Localization and characterization of the hyaluronan-binding site on the link module from human TSG-6. Structure. 2000; 8:763–774. [PubMed: 10903951]
 31. Mahoney DJ, Blundell CD, Day AJ. Mapping the hyaluronan-binding site on the link module from human tumor necrosis factor-stimulated gene-6 by site-directed mutagenesis. J Biol Chem. 2001; 276:22764–22771. [PubMed: 11287417]
 32. Blundell CD, Almond A, Mahoney DJ, DeAngelis PL, Campbell ID, Day AJ. Towards a structure for a TSG-6. hyaluronan complex by modeling and NMR spectroscopy: insights into other members of the link module superfamily. J Biol Chem. 2005; 280:18189–18201. [PubMed: 15718240]
 33. Higman VA, Blundell CD, Mahoney DJ, Redfield C, Noble ME, Day AJ. Plasticity of the TSG-6 HA-binding loop and mobility in the TSG-6-HA complex revealed by NMR and X-ray crystallography. J Mol Biol. 2007; 371:669–684. [PubMed: 17585936]
 34. Higman VA, Briggs DC, Mahoney DJ, Blundell CD, Sattelle BM, Dyer DP, Green DE, DeAngelis PL, Almond A, Milner CM, Day AJ. A refined model for the TSG-6 link module in complex with hyaluronan: use of defined oligosaccharides to probe structure and function. J Biol Chem. 2014; 289:5619–5634. [PubMed: 24403066]
 35. Kohda D, Morton CJ, Parkar AA, Hatanaka H, Inagaki FM, Campbell ID, Day AJ. Solution structure of the link module: a hyaluronan-binding domain involved in extracellular matrix stability and cell migration. Cell. 1996; 86:767–775. [PubMed: 8797823]

36. Blundell CD, Mahoney DJ, Almond A, DeAngelis PL, Kahmann JD, Teriete P, Pickford AR, Campbell ID, Day AJ. The link module from ovulation- and inflammation-associated protein TSG-6 changes conformation on hyaluronan binding. *J Biol Chem.* 2003; 278:49261–49270. [PubMed: 12972412]
37. Parkar AA, Day AJ. Overlapping sites on the Link module of human TSG-6 mediate binding to hyaluronan and chondroitin-4-sulphate. *FEBS letters.* 1997; 410:413–417. [PubMed: 9237673]
38. Wisniewski HG, Snitkin ES, Mindrescu C, Sweet MH, Vilcek J. TSG-6 protein binding to glycosaminoglycans: formation of stable complexes with hyaluronan and binding to chondroitin sulfates. *J Biol Chem.* 2005; 280:14476–14484. [PubMed: 15699048]
39. Heng BC, Gribbon PM, Day AJ, Hardingham TE. Hyaluronan Binding to Link Module of TSG-6 and to G1 Domain of Aggrecan Is Differently Regulated by pH. *Journal of Biological Chemistry.* 2008; 283:32294–32301. [PubMed: 18806261]
40. Williamson MP. Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy.* 2013; 73:1–16. [PubMed: 23962882]
41. Losonczi JA, Andrec M, Fischer MW, Prestegard JH. Order matrix analysis of residual dipolar couplings using singular value decomposition. *Journal of magnetic resonance.* 1999; 138:334–342. [PubMed: 10341140]
42. Prestegard JH, Bougault CM, Kishore AI. Residual dipolar couplings in structure determination of biomolecules. *Chemical reviews.* 2004; 104:3519–3540. [PubMed: 15303825]
43. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc.* 2003; 125:1731–1737. [PubMed: 12580598]
44. de Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin AM. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets, 0 on the CAPRI targets. *Proteins.* 2007; 69:726–733. [PubMed: 17803234]
45. Day AJ, Aplin RT, Willis AC. Overexpression, purification, and refolding of link module from human TSG-6 in *Escherichia coli*: effect of temperature, media, and mutagenesis on lysine misincorporation at arginine AGA codons. *Protein expression and purification.* 1996; 8:1–16. [PubMed: 8812829]
46. Kahmann JD, Koruth R, Day AJ. Method for quantitative refolding of the link module from human TSG-6. *Protein expression and purification.* 1997; 9:315–318. [PubMed: 9126602]
47. Pomin VH, Park Y, Huang RR, Heiss C, Sharp JS, Azadi P, Prestegard JH. Exploiting enzyme specificities in digestions of chondroitin sulfates A and C: Production of well-defined hexasaccharides. *Glycobiology.* 2012; 22:826–838. [PubMed: 22345629]
48. Kneller DG, Kuntz ID. UCSF Sparky - and NMR display, annotation and assignment tool. *Journal of Cellular Biochemistry.* 1993:254–254.
49. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR.* 1995; 6:277–293. [PubMed: 8520220]
50. Farmer BT 2nd, Constantine KL, Goldfarb V, Friedrichs MS, Wittekind M, Yanchunas J Jr, Robertson JG, Mueller L. Localizing the NADP⁺ binding site on the MurB enzyme by NMR. *Nature structural biology.* 1996; 3:995–997. [PubMed: 8946851]
51. Liu Y, Prestegard JH. A device for the measurement of residual chemical shift anisotropy and residual dipolar coupling in soluble and membrane-associated proteins. *Journal of biomolecular NMR.* 2010; 47:249–258. [PubMed: 20506033]
52. Yao L, Ying J, Bax A. Improved accuracy of ¹⁵N-¹H scalar and residual dipolar couplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins. *Journal of biomolecular NMR.* 2009; 43:161–170. [PubMed: 19205898]
53. Kay LE, Ikura M, Tschudin R, Bax A. 3-Dimensional Triple-Resonance Nmr-Spectroscopy of Isotopically Enriched Proteins. *Journal of magnetic resonance.* 1990; 89:496–514.
54. Muhandiram DR, Kay LE. Gradient-Enhanced Triple-Resonance 3-Dimensional Nmr Experiments with Improved Sensitivity. *J Magn Reson Ser B.* 1994; 103:203–216.
55. Zuiderweg ER, Fesik SW. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry.* 1989; 28:2387–2391. [PubMed: 2730871]

56. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. Overcoming the overlap problem in the assignment of ¹H NMR spectra of larger proteins by use of three-dimensional heteronuclear ¹H-¹⁵N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry*. 1989; 28:6150–6156. [PubMed: 2675964]
57. Vistica J, Dam J, Balbo A, Yikilmaz E, Mariuzza RA, Rouault TA, Schuck P. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem*. 2004; 326:234–256. [PubMed: 15003564]
58. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of biomolecular NMR*. 1999; 13:289–302. [PubMed: 10212987]
59. Guntert P, Mumenthaler C, Wuthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol*. 1997; 273:283–298. [PubMed: 9367762]
60. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of biomolecular NMR*. 2002; 24:171–189. [PubMed: 12522306]
61. Valafar H, Prestegard JH. REDCAT: a residual dipolar coupling analysis tool. *Journal of magnetic resonance*. 2004; 167:228–241. [PubMed: 15040978]
62. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *Journal of magnetic resonance*. 2003; 160:65–73. [PubMed: 12565051]
63. De Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*. 2010; 5:883–897. [PubMed: 20431534]
64. Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J Comput Chem*. 2008; 29:622–655. [PubMed: 17849372]
65. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25:1605–1612. [PubMed: 15264254]
66. Kveder M, Pifat G, Pecar S, Schara M, Ramos P, Esterbauer H. Nitroxide reduction with ascorbic acid in spin labeled human plasma LDL and VLDL. *Chemistry and physics of lipids*. 1997; 85:1–12. [PubMed: 9032943]
67. Johnson PE, Brun E, MacKenzie LF, Withers SG, McIntosh LP. The cellulose-binding domains from *Cellulomonas fimi* beta-1, 4-glucanase CenC bind nitroxide spin-labeled cellobiosaccharides in multiple orientations. *J Mol Biol*. 1999; 287:609–625. [PubMed: 10092463]
68. Battiste JL, Wagner G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry*. 2000; 39:5355–5365. [PubMed: 10820006]
69. Jain NU, Venot A, Umemoto K, Leffler H, Prestegard JH. Distance mapping of protein-binding sites using spin-labeled oligosaccharide ligands. *Protein science : a publication of the Protein Society*. 2001; 10:2393–2400. [PubMed: 11604544]
70. Kay LE, Torchia DA, Bax A. BACKBONE DYNAMICS OF PROTEINS AS STUDIED BY N-15 INVERSE DETECTED HETERONUCLEAR NMR-SPECTROSCOPY - APPLICATION TO STAPHYLOCOCCAL NUCLEASE. *Biochemistry*. 1989; 28:8972–8979. [PubMed: 2690953]
71. Blundell, CD.; Kahmann, JD.; Perczel, A.; Mahoney, DJ.; Cordell, MR.; Teriete, P.; Campbell, ID.; Day, AJ. Getting to grips with HA-protein interactions. In: Kennedy, JF.; Phillips, GO.; Williams, PA.; Hascall, VC., editors. *Hyaluronan*. Woodhead Publishing Ltd; Cambridge, UK: 2002. p. 161-172.
72. Blundell CD, Mahoney DJ, Cordell MR, Almond A, Kahmann JD, Perczel A, Taylor JD, Campbell ID, Day AJ. Determining the molecular basis for the pH-dependent interaction between the link module of human TSG-6 and hyaluronan. *Journal of Biological Chemistry*. 2007; 282:12976–12988. [PubMed: 17307731]

73. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR*. 2009; 44:213–223. [PubMed: 19548092]
74. Wang X, Bansal S, Jiang M, Prestegard JH. RDC-assisted modeling of symmetric protein homooligomers. *Protein science : a publication of the Protein Society*. 2008; 17:899–907. [PubMed: 18436958]
75. Nivedha AK, Makeneni S, Foley BL, Tessier MB, Woods RJ. Importance of Ligand Conformational Energies in Carbohydrate Docking: Sorting the Wheat from the Chaff. *Journal of Computational Chemistry*. 2014; 35:526–539. [PubMed: 24375430]
76. Banerji S, Wright AJ, Noble M, Mahoney DJ, Campbell ID, Day AJ, Jackson DG. Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nature Structural & Molecular Biology*. 2007; 14:234–239.
77. Bathe M, Rutledge GC, Grodzinsky AJ, Tidor B. A coarse-grained molecular model for glycosaminoglycans: Application to chondroitin, chondroitin sulfate, and hyaluronic acid. *Biophysical Journal*. 2005; 88:3870–3887. [PubMed: 15805173]
78. Sattelle BM, Shakeri J, Almond A. Does Microsecond Sugar Ring Flexing Encode 3D-Shape and Bioactivity in the Heparanome? *Biomacromolecules*. 2013; 14:1149–1159. [PubMed: 23439078]
79. Briggs D, Birchenough H, Tariq A, Rugg M, Waltho J, Ievoli E, Jowitt T, Enghild J, Richter R, Salustri A, Milner C, Day A. Metal Ion-dependent Heavy Chain Transfer Activity of TSG-6 Mediates Assembly of the Cumulus-Oocyte Matrix. *J Biol Chem*. 2015; 290:28708–28723. [PubMed: 26468290]
80. Discher DE, Janmey P, Wang YL. Tissue cells feel and respond to the stiffness of their substrate. *Science*. 2005; 310:1139–1143. [PubMed: 16293750]
81. Discher DE, Mooney DJ, Zandstra PW. Growth Factors, Matrices, and Forces Combine and Control Stem Cells. *Science*. 2009; 324:1673–1677. [PubMed: 19556500]
82. Geiger B, Spatz JP, Bershadsky AD. Environmental sensing through focal adhesions. *Nature Reviews Molecular Cell Biology*. 2009; 10:21–33. [PubMed: 19197329]
83. Deshauer C, Morgan AM, Ryan EO, Handel TM, Prestegard JH, Wang X. Interactions of the Chemokine CCL5/RANTES with Medium-Sized Chondroitin Sulfate Ligands. *Structure*. 2015; 23:1066–1077. [PubMed: 25982530]
84. Dyer DP, Thomson JM, Hermant A, Jowitt TA, Handel TM, Proudfoot AEI, Day AJ, Milner CM. TSG-6 Inhibits Neutrophil Migration via Direct Interaction with the Chemokine CXCL8. *Journal of Immunology*. 2014; 192:2177–2185.
85. Sanggaard KW, Sonne-Schmidt CS, Jacobsen C, Thogersen IB, Valnickova Z, Wisniewski HG, Enghild JJ. Evidence for a two-step mechanism involved in the formation of covalent HC center dot TSG-6 complexes. *Biochemistry*. 2006; 45:7661–7668. [PubMed: 16768462]

Figure 1.

Superimposed ^1H - ^{15}N HSQC spectra of Link_TSG6 titrated with C444S at a constant concentration of protein: black-control, without ligand; blue-40% ligand; cyan-80%; green-100%; magenta-150%; red-200%. Solid arrows represent residues showing fast exchange and circles show resonances experiencing slow exchange in the presence of a sugar bond.

Figure 2.

Chemical shift perturbation ($\Delta\delta$) of Link module caused by C444S relative to free protein. A) Chart shows the absolute chemical shift difference in the presence of C444S compared with free protein. Red and blue indicate residues belonging to fast and slow exchange regimes, respectively. Residues with peaks that disappear upon ligand binding are indicated by an open bar marked with an asterisk. B) Mapping of shift perturbations caused by C444S binding onto the Link_TSG6 in its HA bound state (1O7C.pdb). The backbone structure is shown with a ribbon in dark gray. The red and blue colors indicate fast and slow exchange as in 2A.

Figure 3.

Fitting for intensity of residues in slow exchange (A). Fitting of chemical shift for residues in fast exchange (B). In A, averaged normalized intensities of residues, A48, G58, and G74 were used for uncomplexed Link_TSG6 (open circle with dotted fitting line) and complexed Link_TSG6 (closed circle with solid fitting line). For B normalized chemical shift change for residues W51, A53, and Y91 were used (closed diamond with solid fitting line). The fitting line was generated using equation 1 and the solution to equations 2-4 for slow and fast exchange, respectively.

Figure 4.

Expanded ^1H - ^{15}N HSQC spectra of Link_TSG6 in the presence of TEMPO analogs. Comparison of ^1H - ^{15}N HSQC spectra of Link_TSG6 with two-fold molar excess of C444S-TEMPO oxidized (A), and reduced (B). Solid red circles and red labels indicate peaks reappearing on reduction of nitroxide, along with their assignment, Dotted red circles show the corresponding peak positions in the presence of oxidized TEMPO.

Figure 5.

R_1 , R_2 , and R_2/R_1 data for free (Δ) and C444S-bound ($\bullet$) Link_TSG6 at 800MHz, 25 °C. All samples were dissolved in 50 mM MES, pH 6.0.

Figure 6. Superimposition of backbone ribbons for 20 solution structures of Link_TSG6 in the presence of C444S hexasaccharide. The fold is composed of two small anti-parallel β -sheets (blue) and two α -helices (red).

Figure 7. Models for the dimer of TSG-6 Link module in complex with C444S-TEMPO. 10 top-scored models are shown which are classified into two groups, A and B. Group A shows 5 models with ligand binding to a surface comprised of residues 62–76, which does not correlate well with the shift mapping data (see Figure 2). Group B shows 5 models with ligand docking to the HA binding groove and composed of the slow exchange residues from the titration experiment. The protein backbone is shown in dark grey.

Figure 8.

Lowest energy model of the Link_TSG6 dimer in complex with C444S-TEMPO. (A) shows the dimer model with a ribbon and surface structure (shown in gray and blue for each monomer). The two docked ligands are shown as stick structures (sulfur atoms colored yellow) with surfaces shown in magenta and cyan. Link_TSG6 residues showing fast exchange effects on C444S binding, which correspond to the second weaker binding site, are colored in red, orange and yellow. The surface for residues with the largest perturbation (>0.07 ppm) is indicated in red, moderate perturbations (>0.05 ppm) are indicated in orange, and small perturbations (>0.02 ppm) are indicated in yellow (B) shows an expansion of binding interface between the Link module and C444S-TEMPO. The residues in close contact to sugar are shown with the labels.

Figure 9. Mapping of HA binding sites and slow exchange shift perturbations caused by CS binding to the Link_TSG6 structure (1O7C.pdb). A) Residues involved in HA binding are indicated in cyan. B) CS perturbed residues are depicted in red (slow exchanging and disappearing residues) and orange (intermediate exchanging residue).

Figure 10.

Superimposition of structures of Link_TSG6 in its C444S and HA bound states. Backbone structures of HA (1O7C.pdb, pink ribbon) and C444S (blue ribbon) are shown with the N- and C- termini labeled. The inset has some of the putative C444S binding residues labeled.

Table 1

Peak intensity and derived amide proton to nitroxide distance for backbone resonances experiencing reduction of peak intensities

Resonance	Link_TSG6 (no ligand)	With spin label (oxidized)	With no spin label (reduced)	Distances (Å)
G79	1.00	0.12	0.65 *	12 ± 4
I80	1.00	0.00	0.89 *	<10
N83	1.00	0.19	0.43 *	13 ± 5

* Incomplete nitroxide reduction in the treatment with ascorbate

Table 2Table of sedimentation equilibrium results (at 30,000 rpm) for a 50 μ M Link_TSG6 sample

C444S concentration (μ M)	Log ₁₀ K _A	RMSD [*]	Mass Da (ideal single species)
0	-	0.0043	10533
50	3.04	0.0052	12589
500	3.24	0.0092	13895
1000	3.82	0.0051	14004

* Root mean square deviation around the fit; this is for a single species model in the case of no C444S added, and for a monomer-dimer model in the presence of C444S.

Table 3

NMR and structural statistics for human Link_TSG6. Structural statistics were computed for the ensemble of 20 NMR structures deposited in the Protein Data Bank. r.m.s., root mean square; r.m.s.d., root mean square deviation.

Conformationally restricting constraints^a	
Distance constraints	
Total	1329
Intraresidue ($i = j$)	260
Sequential ($ i - j = 1$)	394
Medium range ($1 < i - j < 5$)	204
Long range ($ i - j > 5$)	471
Distance constraints per residue	13.6
Dihedral angle constraints	135
No. of constraints per residue	15.0
No. of long-range constraints per residue	4.8
Residual constraint violations^a	
Average no. of distance violations per structure	
0.1–0.2 Å	11.45
0.2–0.5 Å	7.55
>0.5 Å	0.35
Average r.m.s. distance violations/constraint (Å)	0.03
Maximum distance violation (Å)	0.88
Average no. of dihedral angle violations per structure	
1–10°	18.9
>10°	0.3
Average r.m.s. dihedral angle violation/constraints (°)	1.49
Maximum dihedral angle violation (°)	16.80
r.m.s.d. from average coordinates (all/ordered)^{a,b}	
Backbone atoms (Å)	1.2/0.8
Heavy atoms (Å)	1.8/1.4
Ramachandran statistics for ordered residues (Richardson MolProbity)^{a,b}	
Most favored regions (%)	87.9
Additionally allowed regions (%)	11.0
Disallowed regions (%)	1.1
Global quality scores (raw/Z-score)^a	
Verify3D	0.37
ProsaII	0.54
PROCHECKK (ϕ – ψ)	–0.60
PROCHECK (all)	–0.42

MolProbity Clash	21.93
RDC statistics^c	
No. of D _{NH} constraints	46
R	0.99
Q _{r.m.s.}	0.013

^aValues were calculated using PSVS Version 1.5. Average distance violations were calculated using the sum over r^{-6}

^bOrdered residue ranges: 2–9, 13–46, 48–57, 60–62, 66–68, 81–94

Residues selected based on dihedral angle order parameter, with $S(\phi)+S(\psi)\geq 1.8$

^cRDC statistics were computed by PALES