Structural Mechanisms Underlying Sequence-Dependent Variations in GAG Affinities of Decorin Binding Protein A, a *Borrelia burgdorferi* Adhesin

Ashli M. Morgan & Xu Wang

Department of Chemistry & Biochemistry, Arizona State University, Tempe, AZ, 85287

**Corresponding author**: Xu Wang, Department of Chemistry & Biochemistry, Arizona State University, Tempe, AZ, 85287. Phone: 480-727-8256. Fax: 480-965-2747. Email: xuwang@asu.edu

**Short title**: Structural Determinants of DBPA-GAG Interactions

**Keywords**: Glycosaminoglycan, glycosaminoglycan-binding protein, adhesin, Lyme disease, NMR

**Abbreviations**: DBPA, decorin-binding protein A. DBPB, decorin-binding protein B. GAG, glycosaminoglycan. DS, dermatan sulfate. HSQC, heteronuclear single quantum coherence spectroscopy.

**Summary statement**: Structural mechanisms behind variations in GAG affinities of DBPAs from different Borrelia strains were investigated using NMR. DBPA from strain PBr was revealed to have an additional GAG-binding epitope and a retracted linker allowing more access to its GAG-binding sites.
Abstract

Decorin binding protein A (DBPA) is an important surface adhesin of the bacterium *Borrelia burgdorferi*, the causative agent of Lyme disease. DBPA facilitates the bacteria’s colonization of human tissue by adhering to glycosaminoglycan (GAG), a sulfated polysaccharide. Interestingly, DBPA sequence variation among different strains of *Borrelia* spirochetes is high, resulting in significant differences in their GAG affinities. However, the structural mechanisms contributing to these differences are unknown. We determined the solution structures of DBPAs from strain N40 of *Borrelia burgdorferi* and strain PBr of *Borrelia garinii*, two DBPA variants whose GAG affinities deviate significantly from strain B31, the best characterized version of DBPA. Our structures revealed that significant differences exist between PBr DBPA and B31/N40 DBPAs. In particular, the C-terminus of PBr DBPA, unlike C-termini from B31 and N40 DBPAs, is positioned away from the GAG-binding pocket, and the linker between helices one and two of PBr DBPA is highly structured and retracted from the GAG-binding pocket. The repositioning of the C-terminus allowed the formation of an extra GAG-binding epitope in PBr DBPA, and the retracted linker give GAG ligands more access to the GAG-binding epitopes than other DBPAs. Characterization of GAG ligands’ interactions with wild type PBr and mutants confirmed the importance of the second major GAG-binding epitope and established the fact that the two epitopes are independent of one another and the new epitope is as important to GAG binding as the traditional epitope.

Introduction

Glycosaminoglycans (GAGs) are extracellular sulfated polysaccharides ubiquitous to all mammals. GAGs have diverse bioactivities and play a vital role in physiological phenomena ranging from blood coagulation and inflammation to cell differentiation and growth. GAGs are also targeted by bacteria and viruses as entry points to cells or as anchors for establishing residency.[1-3] One of these microbes is the bacterium *Borrelia burgdorferi*, the causative agent of Lyme disease. *B. burgdorferi* is an extracellular bacterium that depends on strong associations with components of extracellular matrix for secure attachment to tissues. To accomplish this, *B. burgdorferi* has evolved to have a number of different adhesins targeting the extracellular matrix.[4, 5] Among the most well studied of these is decorin-binding protein A (DBPA). DBPA was first identified as the adhesin responsible for binding to the proteoglycan decorin, an essential component of connective tissues.[6, 7] Later studies showed the GAG portion of decorin is the major binding site of DBPA, although the possibility of the core protein also being part of the binding epitope has not been ruled out. Even though the type of GAG found on decorin is mostly dermatan sulfate (DS), DBPA has also been known to bind other types of GAGs including heparin and heparan sulfate.[8] In fact, DBPA’s affinity for heparin is considerably higher than that for DS [8, 9].

One particularly interesting feature of DBPA is that the sequence variation of DBPA among different strains of *Borrelia* genospecies is large, with pair-wise sequence identities ranging from 60 to 99.5 %.[10, 11] This magnitude of genetic variations is not observed in DBPA’s close homologue, decorin binding protein B (DBPB), which is conserved across *Borrelia* spirochetes. Studies carried out by Leong and co-workers on DBPAs from several strains of *Borrelia* spirochetes indicate the affinities of different versions of DBPA vary greatly
and confer different tissue preferences.\[11, 12\] Their results demonstrated that DBPA from the PBr strain of \textit{B. garinii} binds to decorin and DS with high affinity (\(K_D \approx 0.16 \mu M\) for intact DS) while affinity of B31 DBPA for the same ligands are lower (\(K_D \approx 0.63 \mu M\) for intact DS), and DBPA from strain N40 has the weakest DS affinity among the three (\(K_D \approx 2.58 \mu M\) for intact DS).\[11\] Although the GAG-binding epitopes of B31 and N40 versions of DBPA have been investigated \[13, 14\], very little is known about the GAG-binding epitopes of PBr DBPA and the mechanisms by which it achieves its higher GAG affinities.

Knowing the mechanisms leading to the GAG affinity differences among DBPA variants are not only important to understanding the role of this protein in promoting bacterial adhesion, it may also help in designing DBPA-based tools for combating \textit{Borrelia} bacteria. In particular, because of DBPA’s strong tendency to induce human immune response, it has been considered as a component in vaccines and serological tests for Lyme disease.\[15\] However, the large sequence variation in DBPA may require the inclusion of several versions of DBPA to elicit comprehensive immunity against all strains since sequence variations in DBPA may prevent antibodies against DBPA of one strain from detecting other versions. Understanding of the structural differences and similarities among the different strains will aid in the intelligent selection of representative DBPA variants that can present greatest coverage in these tests. The family of DBPA proteins also provides an excellent opportunity to study diversity in GAG-binding motifs. Understanding of the structure-activity relationships of GAG-binding proteins is still rudimentary. DBPAs share a similar tertiary fold, therefore GAG affinity differences should come mainly from differences in arrangements of GAG-binding residues on this common frame. This allows the effects of spatial arrangement of basic amino acids on the protein’s affinity and specificity for GAGs to be investigated.

Previously, structure-activity relationships of DBPAs from the closely related strains of B31 and 297 have been investigated by us and other groups.\[9, 14, 16\] To further understand the correlation of DBPA sequence variation with differences in GAG affinities, we have determined the solution structures for N40 and PBr strains of DBPA, and analyzed the interactions of these proteins with GAG ligands. The structures of both proteins retain the five helical bundle fold of B31 DBPA, with the N40 DBPA structure being almost identical to that of B31 DBPA. However, the structure of PBr DBPA contains significant differences compared to B31 DBPA. In particular, the flexible linker between helices one and two adopts a helical conformation in PBr DBPA while the same segments in B31 and N40 are mostly unstructured. In addition, the C-terminal tail of PBr DBPA has a different orientation than those found in B31 and N40 DBPAs. The functional consequences of these changes are that, unlike B31 and N40 DBPAs, which contains only a single GAG-binding epitope, PBr DBPA contains at least two independent binding epitopes. While the location of one epitope is similar to the major GAG-binding epitopes found in B31 and N40 versions of the protein, the other epitope in PBr DBPA is located on the opposite side of the helical bundle, outside the canonical binding pocket for DBPA. Quantitative analysis of binding dissociation constants using NMR and other techniques showed that the new epitope has as high an affinity for GAGs as the epitope in the traditional binding pocket, and the two epitopes are completely independent of one another. Our results show that the mechanisms of GAG-binding vary greatly among different DBPA variants, and detailed structural study will be required to fully understand the GAG-binding mechanisms of each.

**Experimental**
Expression and purification of B31, N40, and PBr DBPAs

The open reading frames (ORFs) for the wild-type (WT) mature DBPA strains B31 (residues 24-191), N40 (residues 29-194) and PBr (residues 22-185) were synthesized by GenScript Inc. (Piscataway, NJ). Each ORF was cloned into the pHUE vector which results in a fusion protein of His-tagged ubiquitin at the N-terminus of each DBPA. PBr mutants were created using site-directed mutagenesis from the WT sequence. The forward primers for Site 1 and Site 2 are the following: K79S, 5'-GCCGGAATTTATCCTGAAAGCAGCATAAAGCAATCCAGTGCC-3'; K130S, 5'-CTGGAAAGAAATTGACATCCAGACAGCATCCAGGTACGGTG-3'; K160S, 5'-GCCCAGGCAATGGAAGACAGCTGAAACTGTAAC-3'; K77,81S, 5'-GCCCAAGCAATGGAAGACAGCTGAAACTGTAAC-3'; K173,176S, 5'-CAACATGATGCGCTGAGCAATCTGAGCGAAAAAGCTAAGACCGCG-3'. Incorporation of the correct mutations was confirmed through sequencing.

Each plasmid was transformed into E. coli BL21(DE3) and then the bacteria was grown at 37°C to an OD$_{600}$ of 0.5. Unlabeled protein was grown in normal LB while isotopically labeled protein was grown in M9 medium supplemented with $^{15}$NH$_4$Cl or [${}^{13}$C] glucose. The bacteria were induced with 0.5 mM IPTG and incubated overnight at 30°C. After the cells had been harvested, the resuspended pellet was incubated with 1 mg/mL lysozyme and sonicated to lyse the cells. The fusion protein was extracted from the cleared supernatant via Ni affinity chromatography with a 5 mL HisTrap column (GE Life Sciences) and was eluted using an imidazole gradient of 25 mM to 300 mM at a flow rate of 3 mL/min. The fusion protein was collected and exchanged into 25 mM Tris (pH 8.0), 100 mM NaCl and was digested with USP2 and 1 mM DTT overnight at room temperature. The cleaved DBPA was separated using a 5 mL HisTrap column and was eluted in the flow-through which was then collected, concentrated and exchanged into 20 mM acetate pH 5.0, 150 mM NaCl.

Acquisition and analysis of NMR structural data for N40 and PBr DBPAs

NMR data for N40 and PBr structure determination were collected on Varian Inova 800 MHz, Bruker Ultra-Shield 600 and 850 MHz spectrometers. Many of the pulse sequences were included in the pulse sequence packages BioPack and TopSpin. HNCACB, CBCACONH, HNCOCA, and HNCO experiments were collected for backbone atom assignment. For side chain atom assignments, CCONH, HCCONH, HCH-TOCSY, $^{15}$N-edited TOCSY, and $^{13}$C- and $^{15}$N-edited NOESY-HSQC spectra were collected. Both HN and NC RDCs were collected using J-modulated pulse sequence on both proteins aligned in 6% polyacrylamide gel.[17] Data processing was conducted with NMRPipe and analyzing was done with NMRView. Relaxation and paramagnetic relaxation enhancement experiments were collected on Bruker Ultra-Shield 600 MHz and Varian Inova 800 MHz spectrometers. $T_1$, $T_2$, and steady-state heteronuclear NOE experiments were collected for both WT N40 and PBr with and without 24 molar equivalents of heparin dp6. NMRPipe[18] was used to process and NMRView[19] was used to analyze the data. The order parameter ($S^2$) was extracted for each protein using relax, model-free software.[20] PRE of TEMPO-tagged heparin dp6 was measured from $^1$H $T_2$ of backbone amide protons.[21] For the N40 PRE experiment, 6 molar equivalents of TEMPO-tagged heparin dp6 was added to 400 µL of 150 µM protein while 2 molar equivalents were added to 400 µL of 300 µM PBr. Contribution of the TEMPO radical was seen in the difference between $^1$H $T_2$ before and after reduction of the radical with 10 molar equivalent of ascorbate.
Structure determination for N40 and PBr DBPAs

Both $^{13}$C and $^{15}$N isotopically edited NOESY-HSQC data were analyzed and assigned, primarily to identify long-range contacts. Backbone dihedral angles of the $\alpha$-helical residues were calculated using TALOS.[22] The NOESY-HSQC peak lists as well as the dihedral angles were used in CYANA to determine each structure.[23] The resulting structure with minimal long-range violations, along with RDCs, was refined in XPLOR-NIH.[24] Surface electrostatic potential of N40 and PBr DBPAs at an ionic strength of 150 mM were calculated with the program APBS.[25]

Production and modification of heparin and heparin fragments

Heparin (Sigma Aldrich) was dialyzed and lyophilized to remove excess salt prior to being digested with heparinase I. Digestion was allowed to continue until depolymerization was approximately 30% complete to give fragments of varying size. These fragments were separated via size exclusion chromatography on a 2.5 cm x 175 cm column (Bio-Rad Biogel P10) with a flow rate of 0.2 mL/min. Each fragment was collected, dialyzed, and lyophilized prior to HPLC verification of fragment size. The fragments were quantified gravimetrically or with a carbazole assay [26]. The reducing end of heparin dp6 (degree of polymerization 6, i.e. hexasaccharide) was modified with 4-amino-TEMPO, a nitroxide radical, through reductive amination.[14] This TEMPO-labeled GAG fragment was used for the PRE study.

Titrations of DBPA variants with heparin dp6

NMR-monitored titration was used to estimate the $K_D$ values for the interaction between heparin dp6 and DBPA variants analyzed in this study. For each PBr titration, 14 molar equivalents of heparin dp6 were added to 400 µL of 150 µM protein in 2 molar equivalent aliquots; whereas the titration of N40 involved the addition of 25 molar equivalents of heparin dp6 to 400 µL of 150 µM protein in 5 molar equivalent aliquots. This was conducted in 150 mM NaCl, 20 mM acetate buffer at pH 5.0 for WT N40, WT PBr as well as PBr Site 1, Site 2a, and Site 2b. The chemical shift changes noted in each $^1$H-$^{15}$N HSQC were normalized into one chemical shift value using the equation, $\delta_H = [\Delta\delta_H^2 + 1.7\Delta\delta_N^2]^{1/2}$, where $\delta_H$ and $\delta_N$ represent the chemical shifts for $^1$H and $^{15}$N, respectively.[27] The $K_D$ binding curves were fitted in xcrvfit (http://www.bionmr.ualberta.ca/bds/software/xcrvfit/) to plot the normalized chemical shift against the heparin dp6:DBPA ratio. $^1$H-$^{15}$N HSQC spectra were collected on a Bruker Ultra-Shield 600 MHz spectrometer.

Gel mobility shift assays for WT B31, N40, PBr and PBr mutants

Heparin dp8 fragments were fluorescently labeled with 0.1M 2-aminoacridone (2-AMAC) according to the method of Lyon et al.[28] The wildtype (WT) DBPA assay was performed by mixing 1 µg of the fluorescently labeled heparin dp6 with 0.5, 1 or 2 molar equivalents of either B31, N40 or PBr DBPA in 50 mM sodium phosphate (pH 6.5), 150 mM NaCl buffer for a total volume of 12.5 µL. The PBr mutant assay was performed by mixing 1 µg of the fluorescently labeled heparin dp8 with 1 or 2 molar equivalents of DBPA in 50 mM sodium phosphate (pH 6.5), 150 mM NaCl buffer for a total volume of 12.5 µL. The control for both assays is protein-free fluorescent heparin dp6 or dp8 fragments. The reaction mixtures were incubated at room temperature for 30 minutes then were run in a 1% agarose gel at 120 V for 15-25 minutes. A UV panel was used to visualize the shifts.[29]
Enzyme-linked immunosorbent assay (ELISA) analysis of heparin binding to WT PBr and PBr mutants

Intact heparin was biotinylated and quantified with the Biotinylation Quantitation kit (Pierce) according to manufacturer’s instructions. Neutravidin-coated microplates (G-Biosciences) were coated with 100 µL of PBS containing 20 µg/mL of biotinylated heparin and were incubated with 2 µg of His-tagged WT PBr and PBr mutants in 100 µL PBS. The results were quantified using 1:2000 Anti-His HRP (Qiagen) and developed with tetramethylbenzidine (TMB) substrate solution. An additional 100 µL of 0.1 M HCl was added to each well to stop the reaction which induced a color change thus allowing the plate to be read at 450 nm.

Competitive heparin-binding ELISA for PBr DBPA variants and biotinylated N40 DBPA

N40 DBPA was biotinylated by reacting carboxyl side chains in N40 DBPA with amino-PEG₂-biotin (Thermo Scientific). Specifically, 1 mL of 26 μM of N40 DBPA was treated with 2.6 mM of EDC and 5.2 mM of amino-PEG₂-biotin for 30 minutes at room temperature. Excess biotin and EDC were then removed by desalting the mixture with a HiTrap Desalting column (GE Healthcare). The competition ELISA was carried out by first incubating commercially available heparin-coated microwell plates (Bioworld) with 100 µL per well of PBS containing 4 µg of biotinylated N40 DBPA and either 0, 2, 4 or 6 µg of PBr DBPA variants for 1 hour. The amount of bound N40 DBPA was then determined by incubating the wells with streptavidin-HRP (Cell Signaling Inc.) for 1 hour, and followed by treatment with TMB reagents for 20 minutes before the absorbance at 450 nm was measured.

Results

Structural differences among DBPAs of strains B31, N40 and PBr

Solution structure of B31 DBPA is comprised of a five-helix bundle with unstructured segments between the first two helices and at the C-terminus. The crystal structure of DBPA from strain 297 adopts an identical fold. Building on these investigations, we have determined the solution structures of DBPA from strain N40 of Borrelia burgdorferi and strain PBr of Borrelia gariini, two variants that have significantly different GAG affinities than B31 DBPA.[11, 12] Structural statistics for these two structures are shown in Table 1. It is evident from these structures that N40 and PBr DBPAs have similar helical topologies and hydrophobic packings as B31 DBPA (Figure 1). Sequence alignment of all three DBPAs indicates that, despite the low sequence identity between PBr and B31 DBPAs, all three DBPAs share remarkably similar secondary structural element arrangements (Figure 2). Analysis of hydrophobic contacts within each structure showed that both N40 and PBr DBPAs are stabilized by the presence of a hydrophobic core formed with hydrophobic amino acids from helices 2, 3, and 5. The same hydrophobic core is also present in B31 DBPA.

Figure 3 shows the superimposition of B31, N40 and PBr DBPA structures. Both N40 and PBr DBPA structures superimposed with the B31 DBPA structure very well. RMSD between helical residues of B31 DBPA and N40 DBPA is 1.9 Å, and RMSD between helical residues of B31 DBPA and PBr DBPA is 2.3 Å. Despite the similarities in the helical regions, structure of PBr DBPA differs significantly from B31 and N40 DBPAs in two respects. Specifically, the linker between helices 1 and 2 of PBr DBPA (residues 56 to 67) adopts a helical conformation while the same region in B31 and N40 DBPA is mostly unstructured (Figure 3). The helical nature of PBr DBPA’s linker was confirmed by the values of backbone dihedral angles of the linker residues predicted using TALOS+ [30] and by the presence of NOEs.
between sequential amide protons, which are reliable indicators of helical conformation. The structural consequence of such a conformation change is that the PBr DBPA linker is more compact than those in B31 & N40 DBPAs. In addition, PBr does not contain the disulfide bond connecting the C-terminus and helix 5, which is present in both B31 and N40 DBPAs. This allows the C-terminus of PBr to be less restricted and to point away from the canonical GAG-binding pocket (Figure 3). It should also be noted that both N40 and PBr are missing the basic amino acid cluster (the BXBB motif) in the flexible linker, which is known to be important for GAG-binding in B31 DBPA. [14]

Because GAG-protein interactions are mediated mostly by electrostatic interactions between sulfate or carboxyl groups on the polysaccharide and basic amino acids on the protein, a surface electrostatic potential map of the protein is often a good way to identify possible GAG-binding epitopes. Electrostatic potential mappings of N40 and PBr DBPAs confirmed the presence of a basic pocket formed by helices 1, 2 and 5 (Figures 4 and 5) of both proteins. This basic patch, which contains three lysine residues (K82, K163 and K170 in B31 DBPA; K85, K166 and K173 in*N*40 DBPA) known to be crucial to GAG binding, has also been identified as the primary site for DBPA-GAG interactions in both B31 and 297 DBPAs.[13, 14, 16] It is notable that the linker in N40 and B31 DBPAs almost entirely obscures this basic pocket. However, the helical nature of the linker in PBr results in a more exposed basic patch than in B31 and N40 DBPAs (Figure 5). Besides the three conserved lysines (K79, K160, and K167) in the canonical binding site (Site 1), the electrostatic potential map of PBr also indicates the possible existence of a second GAG-binding epitope composed of residues K44, K77, K81, K173, and K176 (Site 2) (Figure 5). This site is located on the opposite side of the helices as Site 1, and is not found in B31 or N40 DBPAs.

**GAG’s interaction with WT N40, B31 and PBr DBPAs**

Previous studies by Leong and co-workers have provided comprehensive analyses of the differential binding affinities of the DBPA variants.[11, 12] Specifically, PBr DBPA has the highest affinity for both decorin and DS, while B31 DBPA shows a slightly weaker affinity, and N40 DBPA has the weakest affinity for these ligands. We also carried out a qualitative comparison of DBPA-heparin hexasaccharide (dp6, degree of polymerization 6) interactions among the three strains with a gel mobility shift assay (GMSA), an experiment based on the principle that the migration of heparin dp6 in electrophoresis will be impeded upon binding to DBPA. Figure S1 shows the results of the assay. After the addition of one molar equivalent of protein to fluorescently labeled heparin dp6, both B31 and PBr DBPAs produced shifts in the positions of a significant fraction of heparin dp6 fragments. In contrast, N40 DBPA induced shifts in a much smaller fraction of heparin dp6, indicating GAG affinity of N40 DBPA is lower than that of B31 and PBr DBPAs. These results are consistent with previous studies.[11, 12]

In order to verify the involvement of the proposed GAG-binding epitopes in binding GAGs, WT PBr’s interaction with heparin dp6 was examined via NMR-monitored titration. Although DS is the GAG-type most commonly found on decorin, our choice of GAG ligands was limited by the fact that solution NMR studies cannot be conducted using natural polysaccharides because of the protein aggregation induced by the long polymers. As a result, only short GAG ligands can be used in these titrations. However, the use of short ligands diminishes the avidity effect that protein-GAG interactions rely on to achieve high affinity. Our previous studies have shown that DBPA’s affinity for DS dp6 is very low (K_D > 10 mM), thus not ideal for identifying GAG-binding epitopes. To improve the binding affinity, we chose to use
heparin dp6 as our ligand. We have used these ligands previously [9, 14], and DBPAs’ relative affinities for these ligands correlate well with their affinities for intact DS in our experience.

In the titrations of both N40 and PBr DBPAs with heparin dp6, significant changes in the chemical shifts of amide protons and nitrogens were seen (Figure 6). However, N40 DBPA required more ligands to produce shifts of similar magnitudes as PBr DBPA. Because magnitudes of chemical shift changes are a good indication of involvement in ligand binding, these values are often used to identify residues in the GAG binding sites. Figure 6 shows the residue-specific changes in amide proton and nitrogen chemical shift changes for both N40 and PBr DBPAs. Residues N68, T90 and C191 in N40 DBPA produced the biggest changes in chemical shifts. These residues are found in the linker, helix 2 and the C-terminus, all are around the basic GAG-binding pocket (Figure 6A). Residues E57, L76, K130, L175, E177, and A179 in PBr DBPA showed the most changes in chemical shifts upon the addition of heparin dp6. These residues are found in helix 2, the linker between helices 3 and 4, as well as the C-terminus (Figure 6B). Most of which are located near either Site 1 or Site 2. It should also be noted that none of the residues in the linker regions of PBr showed large changes in their chemical shifts. This is significantly different from what has been observed in B31 DBPA.[9]

Besides identifying possible GAG-binding epitopes, we also obtained the $K_D$s of interaction for these proteins by measuring the magnitudes of chemical shift changes induced by different concentrations of heparin dp6. The $K_D$ for N40 DBPA’s interactions with heparin dp6 was calculated to be greater than 4 mM for most residues with large chemical shift changes (Figure S2). In contrast, $K_D$ of interaction for B31 DBPA and heparin dp6 is ~ 0.5 mM.[14] Heparin dp6 binding curves of PBr DBPA residues differed significantly and did not produce a consensus $K_D$ value. For residues near Site 1 (T64, K130, G133), the binding curve showed the site was not saturated at high heparin dp6 concentrations (Figure S3), indicating the GAG affinity of the site is low. However, another group of residues at the C-terminus (A179 & A182), near Site 2, produced binding curves that showed much higher affinity for GAGs and quick saturation (Figure S3). Fitting binding curves independently and assuming that the concentration of the free ligand is not affected significantly by the presence of the other binding site, which is valid because the ligand is present in much higher excess than the protein throughout the titration, we obtained $K_D$ of ~ 1.2 mM for Site 1 residues, and ~ 0.4 mM for Site 2 residues.

To estimate GAG-binding induced changes in conformational dynamics of the proteins, we used the Model-Free approach [14, 31, 32] to measure the flexibility of the protein backbone in the presence and absence of the GAG-ligands. The Model-Free approach combines NMR observables such as longitudinal relaxation rates, transverse relaxation rates, and steady state heteronuclear NOE of the atom to derive a parameter known as the order parameter, or $S^2$. Values of $S^2$ range from values 0 to 1, with 1 indicating the atom is in a region of protein having rigid conformation and 0 indicating the atom is in a completely flexible segment of the protein. We measured the $S^2$ values of backbone amide nitrogen atoms in the presence and absence of the heparin dp6 ligand for both N40 and PBr DBPAs. No large changes in $S^2$ were observed in either protein, indicating that backbone dynamics remain unaffected by GAG-binding (Figure S4).

Although chemical shift mapping is widely-used to define ligand binding epitopes, other phenomena in the protein, such as ligand-induced conformation changes, can produce artifacts that maybe misleading. To unambiguously define the GAG-binding epitope, we also probed the protein with heparin dp6 ligands functionalized at the reducing end with the nitroxide radical TEMPO.[14] The paramagnetic center of TEMPO generates an inhomogeneous magnetic field that causes NMR signals of nearby atoms to experience larger relaxation rates; this phenomenon
is termed paramagnetic relaxation enhancement (PRE). PRE is distance-dependent so residues closest to the TEMPO tag will experience the greatest increases in relaxation rates. By measuring the paramagnetic relaxation contribution to the transverse relaxation rates of amide protons, the location of the reducing end of the bound ligand can be determined with high sensitivity. Both N40 and PBr DBPAs were probed with TEMPO-labeled heparin dp6, and the PRE effect on each residue was quantified by measuring changes in transverse relaxation rates of amide protons after the TEMPO radical was reduced by ascorbic acid. Due to the weak interactions of N40 DBPA with GAGs, the PRE effect was not significant even at a protein-ligand ratio of 6 (Figure S5). On the other hand, PBr DBPA experienced significant perturbations after addition of only two molar equivalents of the TEMPO-labeled heparin dp6 ligand. Figure 7 shows the residue-specific PRE of each backbone amide proton in PBr. The residues that showed the highest PRE were found near Site 1(I45, K46, A49, and K79) and Site 2 (E73). These residues are concentrated on helices 1 and 2, indicating that the reducing end of heparin was near these helices. In contrast, the same ligand only produced strong PRE effect in the linker residues (N59 and F60) of B31 DBPA. Besides residues in helices one and two of PBr DBPA, A144 in the short linker between helices four and five also showed a large PRE effect that could indicate either the existence of another weak GAG-binding epitope or non-specific interactions.

Interactions of PBr DBPA mutants with GAGs

The structure of PBr DBPA and the ligand-binding perturbation analyses suggest PBr DBPA most likely possesses two GAG-binding sites. Site 1 is located at the same location as the major GAG-binding epitopes of B31 and N40 DBPAs. However, another cluster of basic amino acids on the other side of the helices 2 and 5 as Site 1 can also be found. This cluster is made up of residues K44, K77, K81, K173 and K176. To determine the contribution of each site to GAG binding, mutants of PBr DBPA missing either Site 1 or Site 2 basic amino acid clusters were created. In the Site 1 mutant, K79, K130 and K160, three basic amino acids that were found in the pocket and showed large changes in either chemical shift mapping analysis or PRE perturbation, were mutated to Ser. In Site 2, residues K77, K81, K173 and K176 were mutated to Ser. Although these residues did not show large chemical shift perturbations themselves, residues around them (L76, L175 and E177) did experience either large chemical shift migrations or TEMPO-induced PRE effects. To further dissect the contributions of each basic amino acid pairs in Site 2 to GAG binding, Site 2a (K77,81S) and Site 2b (K173,176S) mutants were also created. GAG-binding affinities of the mutants were evaluated using NMR titrations, GMSA and ELISA. In NMR titrations, each mutant was titrated with heparin dp6 in the same manner as WT PBr to determine the effect of mutations on $K_D$ of binding. Table 2 lists the apparent heparin dp6 interaction $K_D$ of these mutants. As expected, mutations in PBr DBPA’s Site 1 significantly reduced the chemical shift changes of residues around Site 1 (A49, T64 and G133) after addition of heparin dp6 (Figure S6). However, the apparent $K_D$ estimated from residues near Site 2 is still approximately 0.1 mM, similar to WT PBr (Figure S7). Likewise, mutations of residues K77, K81, K173 and K176 in Site 2 to Ser scarcely changed $K_D$ of Site 1 residues, but greatly reduced heparin-induced chemical shift changes of Site 2 residues. In particular, $K_D$ derived from chemical shift changes of T64, K130 and G133, residues found close to Site 1, decreased from $>1$ mM to $~0.5$ mM (Table 2, Figure S7). These observations indicate that both sites are involved in binding GAGs but act independently so that when one site is removed, the other site retains its GAG affinity. To confirm that K77 and K81 acted synergistically with K173 and K176 in binding GAGs, Site 2a mutants were titrated with heparin dp6, and significant decreases in
heparin-induced chemical shift changes observed in C-terminal residues A179 and A182 were observed, and $K_{D}$s estimated using these residues increased from ~0.4 mM to 0.8 mM (Table 2, Figure S7). Mutations of K173 and K176 to Ser (Site 2b mutant) completely eliminated chemical shift changes A179 and A182 (Figure S6).

To estimate the contributions of each epitope to PBr DBPA’s GAG-binding activity, GMSA and ELISA were carried out on WT PBr and the mutants at physiological pH. Heparin dp8 GMSA showed that, after addition of 1 molar equivalent of protein, a significant fraction of heparin dp6 fragments is bound to WT PBr DBPA, whereas very little of the fragments are bound to the Site 1, full Site 2 and Site 2b mutants, but the Site 2a mutant showed reduced, but consistent shifting of heparin dp8 fragments (Figure 8). These results indicate both sites contribute significantly to GAG binding by PBr DBPA. However, residues K173 and K176 of Site 2 are more important in promoting GAG binding than K77 and K81.

The GMSA results are also consistent with the heparin ELISA assays in which surface bound native heparin polysaccharides were probed with His-tagged WT PBr DBPA and mutants. ELISA assays showed that the four mutants bound to heparin with much lower efficiency than WT PBr DBPA (Figure 9). This includes the Site 2a mutant, which showed significant affinity for heparin dp8 in the GMSA. The discrepancy the Site 2a mutant showed in GMSA and ELISA is most likely due to methodology differences between GMSA and ELISA. In particular, the rate of dissociation of the GAG-Site 2a mutant complex may be significantly higher than between GAG and WT PBr such that majority of Site 2a mutants are actually washed off during ELISA. On the other hand, absorbances of wells coated with mutant PBr DBPAs are slightly higher than control, implying that mutations of a single GAG-binding site in PBr DBPA did not completely eliminate PBr DBPA’s affinity for intact heparin.

We also performed competitive heparin-binding ELISA between N40 and PBr DBPAs using biotinylated N40 DBPA as a reporter. Our competition assay shows that all PBr DBPA variants, including mutants with the weakest GAG affinities were able to displace N40 DBPA effectively in the concentration ranges tested (Figure 10). Furthermore, competition assay performed using His-tagged PBr DBPA variants as the reporter revealed that N40 DBPA was not able to significantly reduce the amount of heparin-bound PBr DBPA variants even when the N40 DBPA concentration is twice as high as PBr DBPAs (Figure S8). This shows PBr DBPA’s affinity for heparin is considerably higher than N40 DBPA, such that loss of one of its GAG binding sites still left the protein with higher GAG affinity than WT N40 DBPA.

Discussion

Because of the important role DBPA plays in establishing early *Borrelia* spirochetes’ colonization [33], understanding how sequence variability in the protein influences its specificity and affinity for GAGs is crucial to providing a rational explanation for differing pathogenicities of different strains of *Borrelia* spirochetes. While there have been many structural studies conducted on DBPAs from strains B31 and N40 of *Borrelia burgdorferi*, other strains, such as PBr, have not been well-studied, making it difficult to rationalize variations in their biological activities. Structural analysis and characterization of GAG interactions for N40 and PBr DBPAs carried out here explain how PBr DBPA’s unique structural characteristics allow it to bind GAGs with greater affinity than B31 and N40 DBPAs. In particular, both B31 and N40 DBPA have a disulfide bond that tethers the C-terminal tail to helix 5. In contrast, PBr DBPA is devoid of both cysteine residues required to form this disulfide bond. As a result, its C-terminus is placed away
from Site 1, allowing the formation of the secondary GAG-binding site by placing residues K173 and K176 on helix 5 close to residues K77 and K81 on helix 2. This site could not have formed without the alternate orientation of the C-terminus.

Besides the differing positions of the C-termini, the linker between helices 1 and 2 in PBr DBPA is also more structured than B31 and N40 DBPAs. The significance of this difference lies in the fact that the flexible, unstructured linkers found in B31 and N40 DBPAs are positioned directly above Site 1 and may hinder GAGs from accessing the GAG-binding pocket. PBr DBPA’s more compact linker leaves more of Site 1 exposed, giving GAG ligands more access to Site 1. This observation is compatible with the heparin dp6-TEMPO perturbation data for PBr DBPA, which showed that the reducing end of heparin dp6 interacts with residues on helices 1 and 2, indicating that the fragment was able to enter the pocket. In comparison, linker residues in B31 DBPA were the ones affected most significantly by PRE[14] while there were very little PRE effects on N40 DBPA residues. Finally, the apparent heparin dp6 K_D of PBr’s Site 1, in the absence of Site 2, is much higher than that of N40 DBPA (0.4 mM vs > 4 mM), even though distributions of basic amino acids in both pockets are similar (Figure S9). These data show that the obstruction of the pocket may be an important factor in determining the GAG-affinity of Site 1. This model of DBPA-GAG interaction also explains why the basic amino acid cluster (the BXBB motif) in the linker of B31 DBPA is important to the protein’s affinity for GAGs: it compensates for the reduced access to the binding pocket by providing additional binding epitopes for GAGs, thereby increasing B31 DBPA’s affinity for GAGs. In contrast, N40 is missing both Site 2 as well as the basic amino acid cluster in the linker, resulting in a much lower affinity for GAGs despite the fact that the number of basic amino acids in N40 DBPA is almost identical to B31 DBPA (28 basic amino acids in N40 DBPA vs. 29 basic amino acids in B31). While PBr does not contain the basic amino acid cluster in its linker, PBr DBPA compensates for this by having a second GAG-binding site and a retracted linker that allows more access to Site 1.

In addition, titrations of PBr DBPA mutants with heparin dp6 showed that the two sites are independent of each other. This independence is manifested in the fact that elimination of either site has no effect on GAG affinity of the remaining site. In fact, the apparent K_DS of binding estimated from several residues close to Site 1 decreased significantly after mutations in Site 2, implying Site 1’s affinity for GAGs increased with elimination of Site 2. Interestingly, the magnitudes of decreases were larger than decreases predicted with the assumption that Site 2 only binds a single ligand. Specifically, because the ligand concentration is much higher than the protein concentration throughout the titration, an active Site 2 should not change the concentration of free ligand significantly if Site 2 only binds a single ligand. One possible explanation for the unexpected large changes in Site 1 K_DS is that Site 2 is capable of simultaneously binding more than one ligand, thus can bind a higher amounts of GAG fragments than expected. Its elimination would increase the concentration of free GAG ligand more than expected. Another possibility is that mutations of Site 2 produced significant changes in the structure of PBr DBPA, leading to enhanced ligand affinity in Site 1. However, this hypothesis is not supported by the fact that the ^15N-edited HSQC’s of both Site 2 mutants are very similar to WT PBr DBPA (Figure S6), indicating the structures has not changed significantly in these mutants. Even though an accurate K_D of interaction is not known for Site 1 of PBr DBPA, the dramatic decreases in PBr DBPA mutants’ affinities for heparin in ELISA indicate both sites are important for PBr DBPA under physiological conditions. It also highlights the importance of
multivalent effect to protein-GAG interactions. Multivalency will be especially critical in vivo where native GAG chains are capable of binding multiple epitopes simultaneously.

These structures of PBr, B31 and N40 DBPAs show that despite having similar topologies, subtle changes in sequences can result in significant structural variations that contribute greatly to differences in the adhesion activities of the protein. Characterization of their structures and GAG interactions can therefore provide crucial insights into how strain variations may contribute to the pathogenicity and offer hypotheses for further in vivo experimentation.

**Accession Numbers**

Coordinates of the 10-structure N40 DBPA ensemble are deposited in the PDB data bank with the accession number 2MTC. Coordinates of the 10-structure PBr DBPA ensemble are deposited in the PDB data bank with the accession number 2MTD. Atom chemical shift assignments of N40 DBPA along with the constraints used to calculate the structure are deposited in the BioMagResBank with the accession number 25156. Atom chemical shift assignments of PBr DBPA along with the constraints used to calculate the structure are deposited in the BioMagResBank with the accession number 25157.

**Acknowledgement**

We want to thank Dr. Brian Cherry of ASU Magnetic Resonance Research Center for maintaining the spectrometers used in this study. Funding for this study is provided by National Institute of General Medical Sciences (R00GM088483) and Arizona State University.

**Author Contribution**

AMM carried out the experiments in the study and participated in analysis of the results. XW designed the experiments and participated in analysis of the results.

**Funding**

Funding for this study is provided by National Institute of General Medical Sciences (R00GM088483) and Arizona State University.
Reference


<table>
<thead>
<tr>
<th></th>
<th>N40</th>
<th>PBr</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of NOE-based distance constraints</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1732</td>
<td>1440</td>
</tr>
<tr>
<td>Intraresidue (i = j)</td>
<td>351</td>
<td>398</td>
</tr>
<tr>
<td>Sequential (</td>
<td>i – j</td>
<td>= 1)</td>
</tr>
<tr>
<td>Medium-range (1 &lt;</td>
<td>i – j</td>
<td>&lt; 5)</td>
</tr>
<tr>
<td>Long-range</td>
<td>494</td>
<td>230</td>
</tr>
<tr>
<td>NOE constraints per restrained residue&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10.7</td>
<td>8.8</td>
</tr>
<tr>
<td>No. of dihedral angle constraints</td>
<td>260</td>
<td>280</td>
</tr>
<tr>
<td>Total no. of structures computed</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>No. of structures used</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Residual constraint violations&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of distance violations per structure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1-0.5 Å</td>
<td>32.2</td>
<td>41.3</td>
</tr>
<tr>
<td>&gt;0.5 Å</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No. of dihedral angle violations per structure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-10°</td>
<td>5.6</td>
<td>9.0</td>
</tr>
<tr>
<td>&gt;10°</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMSD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All backbone atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>1.3 Å</td>
<td>2.5 Å</td>
</tr>
<tr>
<td>Ordered&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.7 Å</td>
<td>0.6 Å</td>
</tr>
<tr>
<td>All heavy atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>1.7 Å</td>
<td>2.7 Å</td>
</tr>
<tr>
<td>Ordered&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.3 Å</td>
<td>1.1 Å</td>
</tr>
<tr>
<td>Ramachandran plot summary from Procheck&lt;sup&gt;d&lt;/sup&gt; (%):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>93.2</td>
<td>90.4</td>
</tr>
<tr>
<td>Additionally allowed regions</td>
<td>5.9</td>
<td>9.3</td>
</tr>
<tr>
<td>Generously allowed regions</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Disallowed regions</td>
<td>0.6</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<sup>a</sup> There are 162 residues (for N40) and 163 residues (for PBr) with conformationally restricting constraints.  
<sup>b</sup> Calculated for all constraints for the given residues, using sum over r<sup>6</sup>.  
<sup>c</sup> Residues with sum of phi and psi order parameters > 1.8. Ordered residue ranges for N40 are the following: 35-59,79-104, 112-133,136-148 and 152-183. Ordered residue ranges for PBr are the following: 34-52, 73-100,106-126,130-141 and 147 to 179.  
<sup>d</sup> Residues selected on basis of dihedral angle order parameter, with S(φ)+S(ψ) ≥ 1.8. Selected residue ranges for N40 and PBr are the same as the ordered residue ranges.
Table 2. $K_D$ of mutant PBr-heparin dp6 interaction from calculation using chemical shift changes from residues K130 (Site 1) and A179 (Site 2).

<table>
<thead>
<tr>
<th>PBr DBPA</th>
<th>$K_D$ (mM)</th>
<th>Site 1 (K130)</th>
<th>Site 2 (A179)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>~ 1.2</td>
<td>~ 0.4</td>
<td></td>
</tr>
<tr>
<td>Site 1 Mutant</td>
<td>--</td>
<td>0.1 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Site 2 Mutant</td>
<td>0.5 ± 0.1</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Site 2a Mutant</td>
<td>0.2 ± 0.1</td>
<td>0.79 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Site 2b Mutant</td>
<td>0.4 ± 0.1</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1. Ensemble of the 10 lowest-energy structures of (A) N40 DBPA and (B) PBr DBPA in solution. Helix 1 (residues 33-60 for N40, residues 30-50 for PBr) is colored green. Helix 2 (residues 79-105 for N40, residues 72-99 for PBr) is colored blue. Helix 3 (residues 112-131 for N40, residues 106-126 for PBr) is colored red. Helix 4 (residues 137-145 for N40, residues 131-141 for PBr) is colored cyan. Helix 5 (residues 157-183 for N40, residues 147-174 for PBr) is colored purple. The structured linker (residues 56-67) of PBr is colored yellow, and the disulfide bond anchoring the C-terminus to helix 5 in N40 is colored orange. The schematic topology of N40 is shown at the bottom left and PBr, the bottom right.

Figure 2. Sequence alignment of B31, N40, and PBr DBPAs. The helical regions of each protein are highlighted in black. The structured linker found in PBr DBPA is notated in grey.

Figure 3. Superimposition of B31 (cyan), N40 (green) and PBr (tan) DBPA structures. The linker in PBr DBPA is more retracted than the linkers in B31 and N40 DBPAs. The C-terminus in PBr DBPA is also not tethered to helix 5 by a disulfide bond as it is in B31 and N40 DBPAs.

Figure 4. (A) Electrostatic potential surface map of N40 DBPA. Red indicates acidic regions and blue, basic regions. (B) Ribbon diagram of N40 DBPA showing the protein in the same orientation as Figure 1A.

Figure 5. Electrostatic potential surface maps of PBr DBPA in the same orientation as Figure 1B (A) and rotated 180° about the vertical axis (C). Red indicates acidic regions and blue, basic regions. Ribbon diagrams of PBr DBPA showing the protein in the same orientation as Figure 1B (C) and rotated 180° about the vertical axis (D).

Figure 6. Titrations of WT N40 and PBr DBPAs with heparin dp6. (A) 15N-HSQC overlays of WT N40 DBPA in the presence of increasing concentrations of heparin dp6. The blue contour represents the initial HSQC spectrum of N40 in the absence of heparin dp6. Each subsequent colored contour represents the HSQC spectrum of N40 at the following concentrations of heparin dp6: 0.75, 1.50, 2.25, 3.00, and 3.75 mM. The concentration of N40 was 0.15 mM. Residues experiencing a large migration are indicated with the residue number and the direction of migration. Normalized chemical shift changes for the residues are shown on the left. The secondary structure for each protein is displayed at the top of each plot. (B) Titration of 0.15 mM WT PBr DBPA with heparin dp6. The color scheme is identical to that of N40 DBPA titration except for the following: each subsequent colored contour after the blue contour (absence of heparin dp6) represents the HSQC spectrum of PBr at the following concentrations of heparin dp6: 0.3, 0.6, 0.9, 1.2, 1.5, and 2.1 mM.

Figure 7. (A) Residue-specific PRE on backbone amide protons of WT PBr from TEMPO-labeled heparin dp6. Residues on helices 1 and 2 (K46, A49, E73, and K75) experienced larger
PRE effects when probed with TEMPO-heparin dp6. (B) Ribbon diagram of WT PBr DBPA with the residues experiencing the greatest PRE effect colored in red.

**Figure 8.** GMSA of heparin dp8 in the presence of increasing concentrations of WT PBr DBPA, Site 1 Mutant, Site 2 Mutant, Site 2a Mutant, and Site 2b Mutant.

**Figure 9.** Effect of mutations in PBr DBPA on GAG-binding as determined by ELISA. Mutations in either GAG-binding epitope reduced PBr DBPA’s affinity for GAGs.

**Figure 10.** Effect of PBr DBPA on binding of biotinylated N40 DBPA to heparin coated microwell plates. Each well contains 4 μg of biotinylated N40 DBPA and either 0, 2, 4 or 6 μg of PBr DBPA variants. All PBr variants easily competed off biotinylated N40 DBPA.
Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.
Figure 6.
Figure 7.
Figure 8.
Figure 10.