THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 18, Issue of May 5, pp. 13564–13570, 2000
Printed in U.S.A.

Heparin-binding Growth-associated Molecule Contains Two Heparin-binding β-Sheet Domains That Are Homologous to the Thrombospondin Type I Repeat*

Received for publication, June 17, 1999, and in revised form, January 10, 2000

Ilkka Kilpeläinen‡§¶, Marko Kaksonen,§‖, Tarja Kinnunen‖, Hanna Avikainen‡, Melissa Fath‡‡, Robert J. Linhardt**‡‡, Erkki Raulo‖, and Heikki Rauvala‖**

*From the ‡NMR Laboratory, Institute of Biotechnology, and ‖Laboratory of Molecular Neurobiology, Department of Biosciences, and Institute of Biotechnology, University of Helsinki, Helsinki FIN-00014, Finland and **Division of Medicinal and Natural Products Chemistry and Department of Chemical and Biochemical Engineering, University of Iowa, Iowa City, Iowa 52242*

Heparin-binding growth-associated molecule (HB-GAM) is an extracellular matrix-associated protein implicated in the development and plasticity of neuronal connections of brain. Binding to cell surface heparan sulfate is indispensable for the biological activity of HB-GAM. In the present paper we have studied the structure of recombinant HB-GAM using heteronuclear NMR. These studies show that HB-GAM contains two β-sheet domains connected by a flexible linker. Both of these domains contain three antiparallel β-strands. In addition to this domain structure, HB-GAM contains the N- and C-terminal lysine-rich sequences that lack a detectable structure and appear to form random coils. Studies using CD and NMR spectroscopy suggest that HB-GAM undergoes a conformational change upon binding to heparin, and that the binding occurs primarily to the β-sheet domains of the protein. Search of sequence data bases shows that the β-sheet domains of HB-GAM are homologous to the thrombospondin type I repeat (TSR). Sequence comparisions show that the β-sheet structures found previously in midkine, a protein homologous with HB-GAM, also correspond to the TSR motif. We suggest that the TSR sequence motif found in various extracellular proteins defines a β-sheet structure similar to that found in HB-GAM and midkine. In addition to the apparent structural similarity, a similarity in biological functions is suggested by the occurrence of the TSR sequence motif in a wide variety of proteins that mediate cell-to-extracellular matrix and cell-to-cell interactions, in which the TSR domain mediates specific cell surface binding.

Heparin-binding growth-associated molecule (HB-GAM)[1] (p18) was originally isolated from rat brain as an 18-kDa neurite outgrowth-promoting protein, the expression of which in brain tissue peaks during the perinatal stage of rapid axon growth and synapse formation (1). HB-GAM is highly homologous with the midkine (MK) sequence (2–4), and these proteins thus form a two-member family of small extracellular proteins that are conserved in vertebrates.

In developing tissues HB-GAM associates with extracellular matrix of axonal tracts and of synapses (5, 6). It is also clearly expressed in developing basement membranes outside of brain (7) and in the cartilage matrix (8). N-syndecan (syndecan-3) acts as a receptor of HB-GAM in brain neurons *in vitro* (9) and localizes in many anatomical areas to the same developing fiber tracts as HB-GAM (10, 11). The heparan sulfate structure of brain N-syndecan is exceptionally heparin-like, especially in its high content of 2-*0*-sulfo-iduronic acid residues, which is of importance in the HB-GAM binding carbohydrate epitope, the minimal size of which appears to be 10 monosaccharide residues (12). The neurite outgrowth-promoting effect, based on HB-GAM/N-syndecan interaction, was very recently shown to be mediated by the cortactin/src-kinase signaling pathway to the cytoskeleton of neurites (13). These findings have led to the concept that N-syndecan mediates HB-GAM-induced neurite growth (for review, see Ref. 14). Furthermore, very recent studies have revealed a role for HB-GAM and N-syndecan in the regulation of hippocampal long-term potentiation (15), a form of brain plasticity implicated in memory and learning.

HB-GAM consists of 136 amino acid residues. It contains an abundance (24%) of cationic, mainly lysine, residues and 5 intrachain disulfide bonds (2, 16). The HB-GAM sequence is highly conserved across different species: >90% identity is observed among the sequences of chicken, rat, bovine, and human (2, 17, 18). The same sequence as that of HB-GAM was found for the mitogenic and neurite outgrowth-promoting protein designated pleiotrophin (18). The designations osteoblast-specific factor-1 (19), heparin binding neurotrophic factor (20) and heparin affin regulatory peptide (21) also refer to the protein having the same sequence as HB-GAM.

To gain further insight into the function of HB-GAM, we have undertaken the study of the structure of this protein and its interaction with heparin by means of heteronuclear NMR. These studies show that HB-GAM contains two β-sheet domains that bind heparin. A search of data bases using the sequences of the β-sheet domains and of the full-length HB-GAM shows that the β-sheet domains are homologous to the thrombospondin type I repeat (TSR) motif that mediates cell surface binding in a number of extracellular matrix and cell surface proteins.

* This work was supported in part by the Academy of Finland, the Technical Research Center of Finland, and the Sigrid Jusélius Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ These authors contributed equally to this work.

¶ To whom correspondence should be addressed: P. O. Box 56, University of Helsinki, Helsinki FIN-00014, Finland. Tel.: 358-9-7085-9540; Fax: 358-9-7085-9541; E-mail: ilkka.kilpelainen@helsinki.fi.

‡‡ Supported by National Institutes of Health Grant HL-52622.

[1] The abbreviations used are: HB-GAM, heparin-binding growth-associated molecule; MK, midkine; TSR, thrombospondin type I repeat; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser and exchange spectroscopy; TOCSY, total correlation spectroscopy.

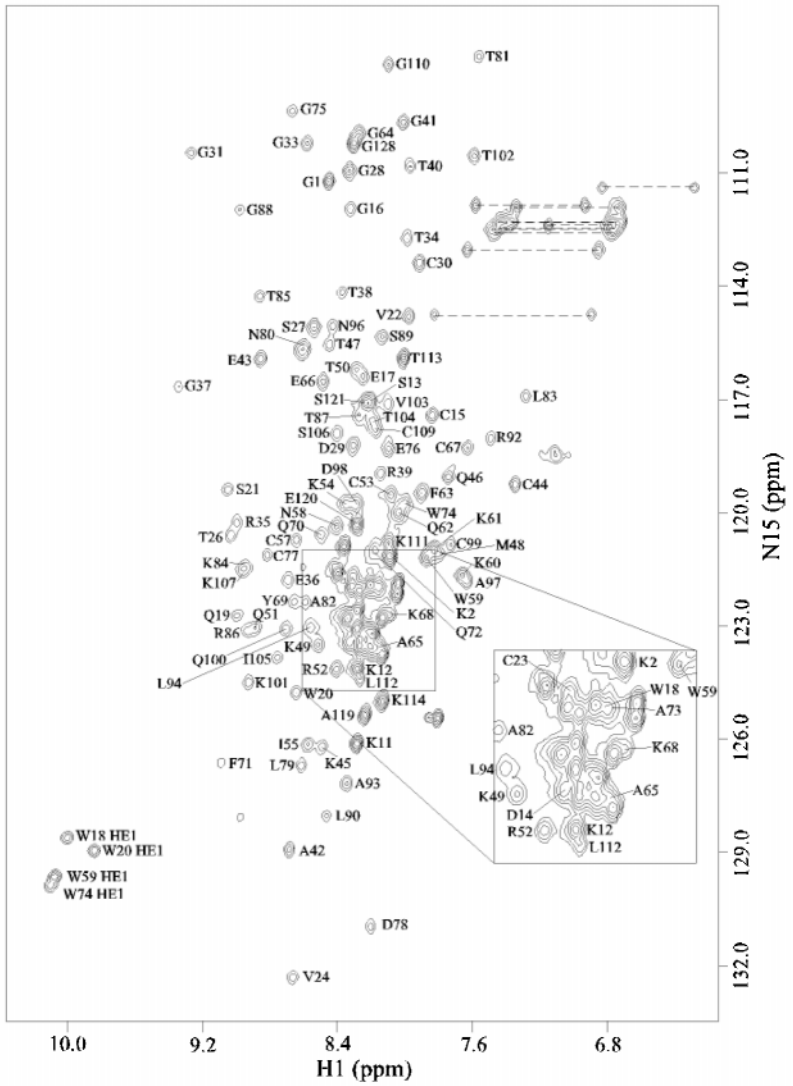

FIG. 1. **^{1}H-^{15}N-HSQC spectrum (750 MHz) of HB-GAM with backbone assignments indicated.** Correlations of side-chain amide groups are shown by *horizontal lines* but have not been assigned. Notably, it was not possible to obtain individual assignment for part of the flexible N- and C-terminal tails of the protein. The amino acid sequence of the HB-GAM produced in *E. coli* is: G(-2)S(-1)GKKEKPEKKVKKSDCGEWQWSVCV-PTSGDCGLGTREGTRTGAECKQTMKT-QRCKIPCNTKKQFGAECKYQF QAWGE-CDLNTALKTRTGSLKRALHNADCQKT-VTISKPCGKLTKPLPQAESKKKKKEGK-KQKEMLD. All the cysteins are in disulfides as follows: Cys^{15}-Cys^{44}, Cys^{23}-Cys^{53}, Cys^{30}-Cys^{57}, Cys^{67}-Cys^{99}, and Cys^{77}-Cys^{109}.

EXPERIMENTAL PROCEDURES

Protein Sample Preparation and Characterization—The baculovirus-derived HB-GAM was used for ^{1}H NMR and has been previously shown to be biologically active and identical to the tissue protein (16). To label HB-GAM with ^{15}N and ^{13}C, we used the glutathione *S*-transferase fusion system J (Amersham Pharmacia Biotech) in *Escherichia coli* (strain BL21(DE3)). The HB-GAM samples were obtained after thrombin cleavage of glutathione *S*-transferase fusion protein. The HB-GAM samples were purified with heparin affinity, glutathione-Sepharose, and ion exchange chromatography. The HB-GAM used in this study contains a Gly-Ser N-domain extension (numbered as -2, -1, respectively). The *E. coli* and baculovirus derived proteins were shown to be essentially identical by matrix-assisted laser desorption and ionization-time of flight mass spectroscopy and ^{1}H NMR (two-dimensional NOESY and TOCSY) spectroscopy. Furthermore, the protein samples were characterized with peptide mapping to ensure that the disulfide bonds were correct in the *E. coli*-derived protein. In this analysis, the protein samples were digested with a proteinase (endoproteinase Lys C or trypsin), and the digested sample was divided into three identical fractions. The first fraction was analyzed with matrix-assisted laser desorption and ionization-time of flight directly; the second one was alkylated and subjected to matrix-assisted laser desorption and ionization-time of flight; and the third was reduced with dithiothreitol, alkylated, and analyzed with matrix-assisted laser desorption and ionization-time of flight mass spectroscopy. The ^{15}N and ^{13}C double-labeled samples were grown in minimal media containing 0.5 g/liter $^{15}NH_4Cl$, 1.25 g/liter glucose, and 1 g/liter Isogro-^{13}C, ^{15}N-Powder growth medium (Isotec). The ^{15}N-labeled samples were produced correspondingly, but without ^{13}C labels. The NMR samples contained 1 mM HB-GAM in 95% $^{1}H_2O$/5% $^{2}H_2O$, pH 4.7.

Expression of Individual HB-GAM Domains in E. coli—Oligonucleotide primers were used to amplify individual β-sheet domains of HB-GAM. The primers for the N-terminal domain (amino acids 45–90) were 5′-AAGGATCCGACTGTGGAGAATGGCAA-3′ and 5′TGGAATTCCTAGTTGCAAGGGATCTTACATCT-3′. For the C-terminal domain (amino acids 97–142) primers 5′-AGGGATCCGCTGAGTGCAAATACCAG-3′ and 5′-TTGAATTCCTAGCCACAGGGCTTGGAGAT-3′ were used. The domain boundaries were determined using the HB-GAM secondary structure NMR coordination map. The PCR products were cloned into the *Bam*HI-*Eco*RI cloning site of pGEX-2T and sequenced. Production of the separate domains was performed in *E. coli* strain BL21. The glutathione *S*-transferase fusion proteins were harvested from bacterial lysates by glutathione-Sepharose and eluted by thrombin cleavage.

Heparin Affinity Chromatography of HB-GAM Domains—The isolated recombinant domains were subjected to heparin affinity chromatography. A linear NaCl gradient from 0 to 2 M in 20 mM sodim phosphate buffer, pH 7.5, was used to compare the elution profiles of separate domains with that of intact HB-GAM produced similarly in *E. coli*.

Preparation of Heparin Tetradecasaccharide—Porcine intestinal mucosal heparin (M_r 14,000) was obtained from Celsus Laboratories (Cincinnati, OH). The heparin (10 g) was digested to 30% completion using heparin lyase I as reported previously (22) and fractioned by Sephadex G50 gel filtration chromatography into sized oligosaccharides. A single tetradecasaccharide (14-mer) was purified from the mixture using strong-anion exchange chromatography, and its purity was assessed as >90% by gradient polyacrylamide gel electrophoresis (23) and capillary electrophoresis (22). One-dimensional ^{1}H NMR and two-dimensional COSY NMR were used to examine the structure of this tetradecasaccharide (22).

NMR Measurements—^{1}H-^{15}N-HSQC, HNCACB, CBCA(CO)NH, HNCO, HCCH-TOCSY, CC(CO)NH, ^{15}N-edited TOCSY and NOESY, and ^{13}C-edited NOESY (24, 25) were acquired with Varian Unity 500 and 600 and Inova 750 spectrometers at 30 °C. Data were processed with $\cos^2$ functions in all dimensions and with linear prediction in ^{15}N and ^{13}C domains in the triple resonance experiments. All spectra were processed on Silicon Graphics O_2 workstations using Felix 97 software (Felix 97.0; Biosym/MSI, San Diego, CA). The ^{15}N T1 and T2 relaxation rates (26) were measured using delays of 0.01, 0.07, 0.14, 0.26, 0.38, 0.53, 0.76, and 1.17 s for T1 and delays of 0.11, 0.14, 0.18, 0.24, 0.32, 0.40, 0.43, 0.48, and 0.58 s for T2. The heteronuclear NOE experiments were recorded with and without 2 s of saturation of the amide protons.

CD Measurements—The binding of heparin (M_r 8000; Sigma) to HB-GAM was followed with circular dichroism (CD). The heparin stock solution was prepared to a 1 mM concentration at pH 6.0. A quartz cuvette with a 0.1-mm optical path was filled with 0.2 mM HB-GAM solution to 200 μl, and the CD spectra were recorded from 250 to 190 nm with a Jasco-720 spectropolarometer.

Sequence Homology Searches—The Gapped-BLAST and the position-specific-iterated BLAST programs (27) were used for searching homologous sequences from the nonredundant protein sequence data base of the National Center for Biotechnology Information. Multiple alignment was constructed using the ClustalX program (28).

RESULTS AND DISCUSSION

NMR Assignments and Domain Structure of HB-GAM—Recombinant HB-GAM purified from the culture medium of Sf9 cells was used to record NMR spectra at 500 and 600 Mhz, because this protein has been previously shown to correspond in its biochemical and cell biological properties to tissue-derived HB-GAM (16). For ^{15}N and ^{13}C labeling, expression systems in bacteria and yeast were explored. Of several expression systems tested, the glutathione *S*-transferase gene fusion system in *E. coli* was found to produce reasonable amounts of HB-GAM that could be purified to apparent homogeneity, as described previously for the baculovirus-derived protein from Sf9 cells (16). NMR spectra and peptide maps (see "Experimental Procedures") recorded from the bacterial recombinant were essentially identical compared with those using the baculovirus-derived protein produced in Sf9 cells, indicating native conformation for the protein produced in *E. coli*. Furthermore, dose-response curves of neurite outgrowth in brain neurons were essentially identical (data not shown) compared with the baculovirus-derived protein studied previously (16). We therefore used the recombinant HB-GAM produced in *E. coli* for further NMR studies.

The 750-MHz ^{1}H-^{15}N-HSQC spectrum of HB-GAM is shown in Fig. 1. A nearly complete backbone assignment was obtained for the structural parts of HB-GAM using triple resonance methods (24, 25, 29, 30). Thus, amide ^{15}N frequencies were obtained from a high-resolution 750-MHz ^{1}H-^{15}N-HSQC spectrum, followed by the collection of corresponding intra- and inter-residual Cα/Cβ connectivities from 500-MHz HNCACB and CBCA(CO)NH spectra. This information allowed in most cases the unique matching of the amino acid residues. The carbonyl shifts were obtained from HNCO. Side-chain proton and carbon chemical shifts were assigned using a combination of three spectra. Partial proton and carbon side-chain assignments obtained from ^{15}N-edited TOCSY and CC(CO)NH were subsequently complemented with those from HCCH-TOCSY. Chemical shifts of all assigned ^{1}H, ^{13}C, and ^{15}N nuclei have been deposited in BioMagResBank.

Hα, Cα, and C′ chemical shift indices (31, 32) indicate that HB-GAM is composed of six β-strands; the consensus crystal

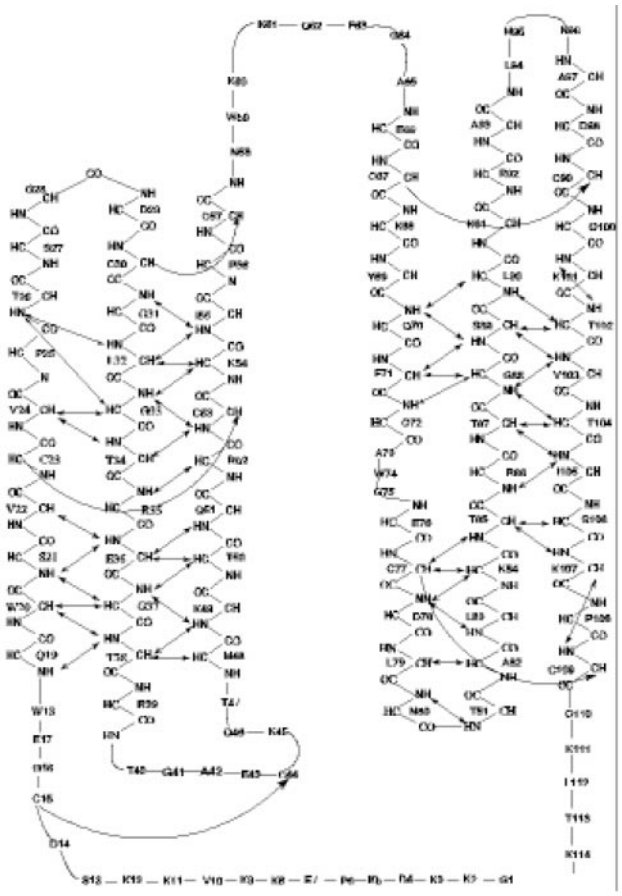

FIG. 2. **Secondary structure of HB-GAM.** *Curved arrows* show the location of disulfide bonds. *Straight arrows* indicate long-range NOE connectivities.

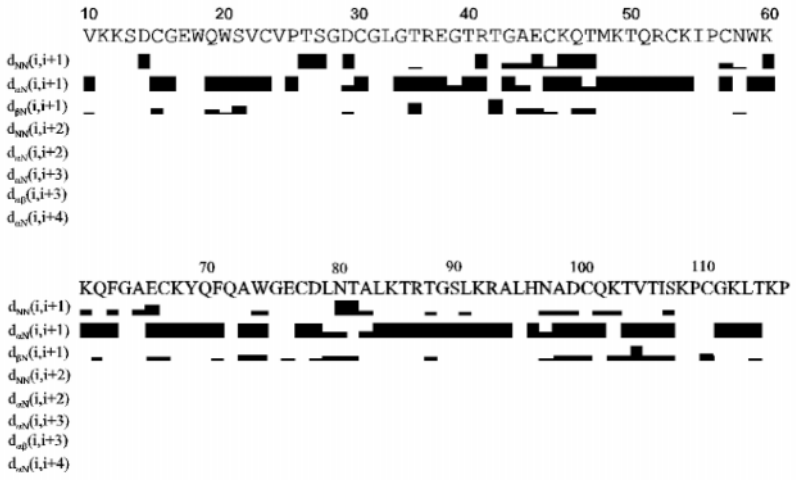

FIG. 3. **Sequential NOE data of HB-GAM.** The *lines* are scaled to the NOE intensities; *bold lines* represent strong NOE connectivities, and *thin lines* indicate weak connectivities.

structure image predicts β-strands for residues 18–27, 32–39, 48–57, 70–78, 83–91, and 101–110 (data not shown). The long-range NOE information (Fig. 2, obtained from ^{15}N- and ^{13}C-edited NOESY at 750 MHz) of HB-GAM in between the β-strands reveals a two-domain structure, similar to that suggested for MK (33). Sequential NOE data are presented in Fig. 3. Both domains consist of three antiparallel β-sheet structures. Heteronuclear-edited NOE data and dipolar coupling data recorded from a partially oriented HB-GAM in a liquid crystalline medium with pf1 phages (34) suggest that the two domains are relatively independent and free to move with respect to each other in solution (data not shown). The linker region between the β-sheet domains of HB-GAM thus appears flexible, jeopardizing attempts to get a high-resolution struc-

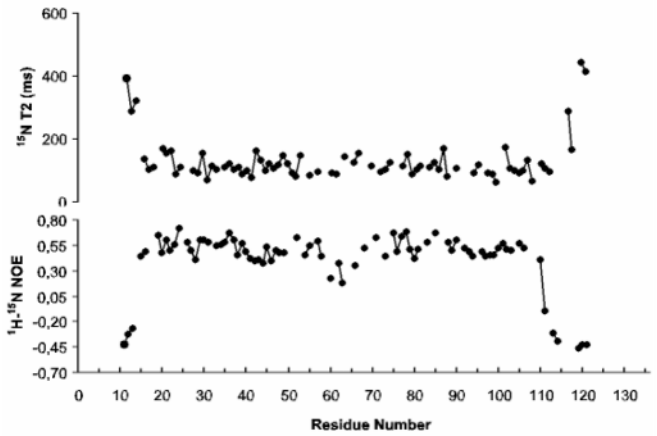

FIG. 4. *Top*, ^{15}N T2 relaxation rates; *bottom*, heteronuclear ^{1}H and ^{15}N NOE intensities plotted against residue numbers.

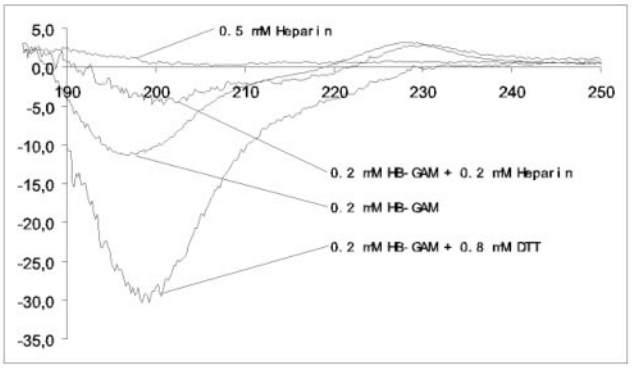

FIG. 5. **Heparin-induced changes in HB-GAM structure as revealed by CD spectroscopy.** CD spectra of 0.2 mM HB-GAM, 0.5 mM heparin, and 0.2 mM 1:1 complex (as molar ratios) of HB-GAM·heparin are shown. Changes in the CD spectra are especially prominent at wavelenghts of 200 and 230 nm.

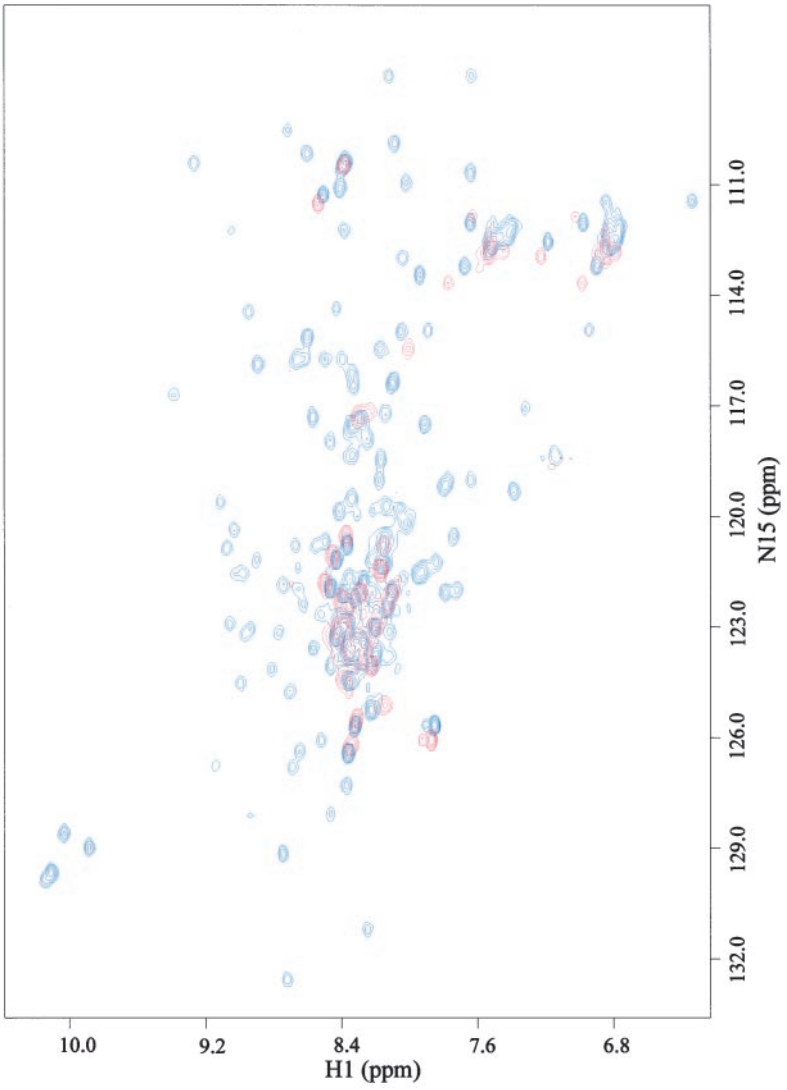

FIG. 6. **Binding of heparin to HB-GAM as revealed by NMR spectroscopy.** Titration of HB-GAM with heparin-derived 14-mer was followed by ^{1}H-^{15}N-HSQC NMR. The spectrum of free HB-GAM is shown in *blue,* and the spectrum of the titration end point is shown in *red* (ratio of HB-GAM/heparin, 1:2). The intense correlations from the flexible, lysine-rich, N- and C-terminal tails in the middle area of the spectra do not change their positions and retain the intense, narrow correlations also in the complex form. The correlations originating from the β-sheet domain areas of HB-GAM have broadened below detection, indicating that both domains take part in the complex formation.

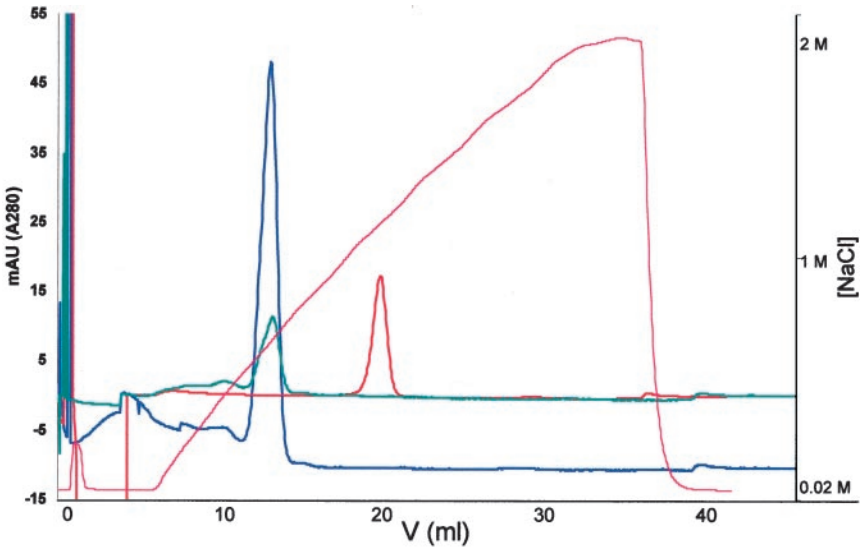

FIG. 7. **Heparin affinity chromatography of intact HB-GAM and separate β-sheet domains.** A heparin-Sepharose HiTrap 1-ml column (Amersham Pharmacia Biotech) was equilibrated to 20 mM phosphate buffer, pH 7.5. A gradient elution program was used to compare the elution profiles of intact HB-GAM (*red*), N-terminal β-sheet domain (*blue*), and C-terminal β-sheet domain (*green*). The NaCl gradient for each run was monitored as conductivity (*magenta*). The intact HB-GAM and its β-sheet domains elute at 1.2 and 0.6 M NaCl, respectively.

ture of the whole molecule under standard NMR conditions.

The N-terminal (amino acids 1–14) and C-terminal (amino acids 111–136) areas of HB-GAM are highly flexible, as indicated by the long T2 ^{15}N relaxation rates and by the heteronuclear NOE intensities presented in Fig. 4. For some parts of the flexible N and C termini it was not possible to obtain individual assignments because of strong overlap in ^{1}H, ^{15}N, and ^{13}C frequencies (Fig. 1). Also, the linker (amino acids 59–66) shows some flexibility, as indicated by slightly longer ^{15}N T2 and smaller NOE values.

Heparin Induces Structural Changes in HB-GAM by Binding to the β-Sheet Domains of the Protein—HB-GAM was initially isolated from crude extracts of brain by heparin affinity chromatography and shown to be displaced from heparin at 1 M NaCl in salt gradient elution, suggesting a strong binding (1). We have previously shown by microtitration calorimetry that the stoichiometry of HB-GAM/heparin interaction is 3 mol of HB-GAM/mol of heparin with a dissociation constant of 460 nM in solution (35). Binding of HB-GAM to heparin was also evident from CD spectroscopy of heparin, HB-GAM, or HB-GAM plus heparin in solution (Fig. 5). Furthermore, CD spectroscopy suggests that HB-GAM undergoes structural changes upon binding to heparin (Fig. 5).

Because the binding of HB-GAM to heparin is relatively tight, and a high molecular weight complex is simultaneously formed, we were not able to follow the movements of correlations in the ^{1}H-^{15}N-HSQC spectrum of HB-GAM during titration with the heparin-derived 14-mer. The 500-MHz ^{1}H-^{15}N-HSQC spectrum of the end point of the titration (at the HB-GAM/heparin molar ratio of 1:2) overlaid with the corresponding spectrum of HB-GAM alone is shown in Fig. 6. The comparison of the ^{1}H-^{15}N-HSQC spectra of HB-GAM with the HB-GAM·heparin complex reveals that the lysine-rich N- and C-terminal tails (although not fully assigned) are not involved in the heparin binding and remain unstructured (Fig. 6) also in the complex form. The signals from the fexible N- and C-terminal tails remain intense and narrow also in the complex form, despite the excess of heparin used and the high molecular weight of the complex. In contrast, the correlations from the β-sheet domains of HB-GAM have broadened significantly, most of them below the detection limit in the complex form (Fig. 6). Thus, although the lysine-rich tails are highly charged, they do not contribute to binding of HB-GAM to heparin, but the binding depends on the β-sheet domains. Based on NMR spectra, both of the β-sheet domains of HB-GAM bind to heparin in solution (Fig. 6). Furthermore, the two β-sheet domains

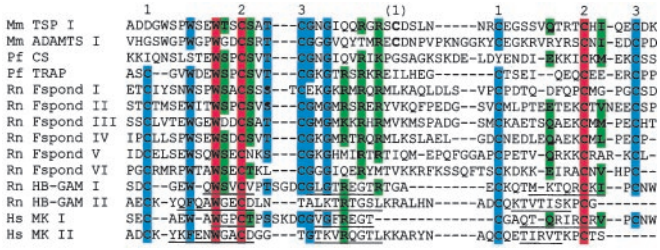

FIG. 8. **Sequences corresponding to the β-sheet structures of HB-GAM and MK are homologous with the TSR sequence motif found in extracellular matrix and cell surface proteins.** Alignment of TSR domains of thrombospondin-1 (*TSP*), ADAMTS-1, CS and TRAP, F-spondin, HB-GAM, and MK is shown. *Roman numerals* after the names indicate the number of the TSR domain in each protein. Residues that are identical in all sequences are shown in *red*; residues that are identical or similar in 80 or 60% of the sequences are shown in *blue* or *green*, respectively. Sequences corresponding to the β-sheet domains of HB-GAM and MK are *underlined*. *Numbers* at the *top* indicate cysteines that are linked by disulfide bonds in HB-GAM and MK. The same disulfide arrangement is also predicted for the other TSR domains (36). In the first two sequences the predicted arrangement of disulfide bonds is slightly different and is indicated by a *number* in *parentheses*. GenBank gene identification numbers are as follows: Mm TSP, 549134; Mm ADAMTS-1, 2809057; Pf CS, 294119; Pf TRAP, 136153; Rn F-spondin (*Fspond*), 544353; Rn HB-GAM, 131554; and Hs MK, 127116. *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Pf*, *Plasmodium falciparum*; *Rn*, *Rattus norvegicus*.

of HB-GAM were expressed as separate recombinant proteins in *E. coli* and were shown to elute at a 0.6 M NaCl concentration in heparin affinity chromatography with salt gradient elution (Fig. 7). Both β-sheet domains of HB-GAM thus display a strong binding to heparin without the polylysine type tails.

Reduction of the disulfide bonds of HB-GAM dramatically reduced heparin binding, in agreement with the NMR results suggesting that native HB-GAM structure is essential for the binding. Thus, whereas the native HB-GAM was displaced from heparin by 1 M NaCl when analyzed by salt gradient elution (1), the reduced HB-GAM only displayed weak binding and was eluted at low salt (~0.2 M NaCl). As expected, NMR and CD spectroscopy (Fig. 5) suggest that the β-sheet structures of HB-GAM are partially destroyed by reduction of the intrachain disulfide bonds.

The β-Sheet Domains of HB-GAM Are Homologous with the TSR Motif—The National Center for Biotechnology Information nonredundant sequence data base was searched with the mature HB-GAM protein sequence using the Gapped-BLAST program (27). The HB-GAM sequence was found to align with

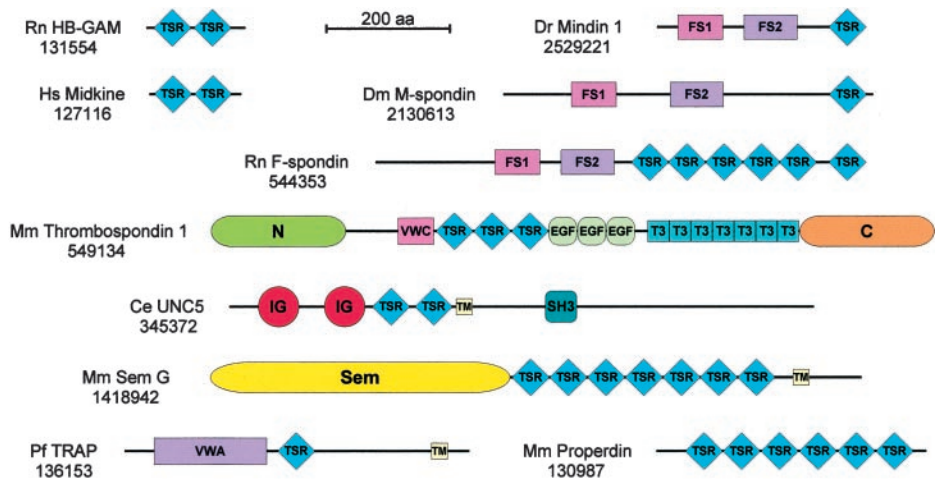

FIG. 9. **Domain structures of selected proteins that contain TSR sequence motifs.** The sequences were retrieved in a position-specific-iterated BLAST search using the HB-GAM sequence. Domain designations: *C*, thrombospondin C-terminal domain; *EGF*, epidermal growth factor-like domain; *FS1* and *FS2*, F-spondin homology 1 and 2; *IG*, Ig domain; *N*, thrombospondin N-terminal domain; *Sem*, semaphorin domain; *SH3*, Src homology 3; *T3*, thrombospondin type III repeat; *TM*, transmembrane region; *TSR*, thrombospondin type I repeat; *VWA*, von Willebrand factor type A domain; *VWC*, von Willebrand factor type C domain. GenBank gene identification numbers are shown under the protein names. *Ce*, *Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*; *Dr*, *Danio rerio*; *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Pf*, *Plasmodium falciparum*; *Rn*, *Rattus norvegicus*.

the TSR repeats in the F-spondin sequence (E value = 0.043). A search using the position-specific-iterated BLAST program (27) gave statistically significant matches with F-spondin and several other proteins that contain TSR domains (Fig. 8). Both β-sheet domains of HB-GAM aligned with with the TSR motif (Fig. 8). The lysine-rich unstructured tails of HB-GAM were not in the alignment.

We then made a position-specific-iterated BLAST query using only the N- or C-terminal β-sheet sequences, and the first matches for both domains after the HB-GAM and MK sequences were the TSR domain sequences (Fig. 8). Search with the MK sequence also gave matches with F-spondin and other proteins having TSR motifs. HB-GAM and MK are thus composed of short N- and C-terminal tails and two TSR domains that are connected with a short linker (Fig. 9).

The main conserved features in the TSR sequences are the cysteines, closely spaced tryptophans, and a cluster of basic residues (Fig. 8). The arrangement of intradomain cysteine bridges in HB-GAM (Fig. 2) and in MK (33) is consistent with that predicted for the TSR motif (Fig. 8 and Ref. 36). Tryptophans are conserved in the HB-GAM and MK N-terminal domains where their spacing is, however, reduced by 1 amino acid. In the C-terminal domains another tryptophan is replaced by phenylalanine.

Our finding that the β-sheet domains of HB-GAM mediate heparin binding agrees with the view that these domains correspond to the TSR sequence motifs, because several TSR repeats have been shown to bind heparin and heparan sulfate. Peptides from thrombospondin TSR bind heparin and inhibit heparin-dependent interaction of cells with thrombospondin and laminin (37, 38). The TSR domains of the malaria proteins CS (39) and TRAP (40) bind to cell surface proteoglycans, an important event in the infection (41, 42). Metalloproteinase ADAMTS-1 uses TSR domains to bind to extracellular matrix heparan sulfates (43).

The conserved tryptophans (Fig. 8) form a heparin binding motif in thrombospondin TSR domains (37). It has been shown that peptides containing a WSXW sequence strongly inhibit heparin binding to thrombospondin (37, 38). Interestingly, reducing the spacing of tryptophan residues did not significantly reduce the inhibitory activity of the peptides (38). A cluster of basic residues downstream of tryptophans has also been shown to contribute to heparin binding in the TSR domain of the malaria CS protein (39). For the midkine C-terminal domain, it has been suggested that the conserved tryptophan in the first β-strand and downstream basic amino acids in the second β-strand are involved in formation of a heparin binding surface (33). The conserved basic residues in the TSR domains are separated by 1 or 3 residues so that on the β-sheet their side chains are on the same side of the structure and thus could participate in forming a heparin binding site (Fig. 8). In an antiparallel β-sheet arrangement, similar to one seen in MK and HB-GAM, they would also be close to the conserved tryptophans. It appears that the TSR domains of different proteins share structural features explaining their heparin binding, but further charactization of the mode of binding is warranted.

Concluding Remarks—The present results show that HB-GAM contains two β-sheet domains, the sequences of which correspond to the TSR motif. Furthermore, our sequence analyses show that the β-sheet domains found recently in the homologous MK protein (33) also correspond to the TSR motif. The amino acid sequences of these four postulated TSR domains found in HB-GAM and MK are quite variable but retain the conserved cysteine/tryptophan motif. The finding that these four polypeptide sequences fold to a clearly similar structure composed of three antiparallel β-strands suggests a general rule for the TSR sequence motif. We are currently producing TSR domains from other proteins to study this hypothesis, because structures of other TSR domains, except for HB-GAM and MK, are not yet known.

Structural data of the TSR domains of HB-GAM and MK also make molecular modeling of other TSR domains a worthy goal. It is noteworthy that many TSR domains are involved in interesting biological interactions (see below), some of which also possess potential for therapeutic applications. This underscores the importance of detailed structural data and molecular modeling studies based on resolved structures.

TSR proteins have functions in specific cell surface interactions. Malaria parasites bind to the surface of hepatocytes and penetrate very rapidly and specifically using their surface proteins CS (44) and TRAP (45). Recent studies suggest that different species of malaria invariably express cell surface proteins having a TSR domain (41). The TSR domain of the malaria sporozoites is used for gliding motility of the parasites, a

prerequisite for infection and a potential target for drug development (41).

It is of interest that several proteins that interact with the neuron surface and are involved in neurite growth and guidance possess TSR domains (Fig. 9). In fact, HB-GAM was initially isolated (1) and cloned (2) as a heparin-binding protein that enhances neurite growth. Other TSR proteins that are implicated in neurite growth and/or guidance include MK (46), UNC5 (36), F-spondin (47), and semaphorins F and G (48). Binding to heparin-type carbohydrates is known to be of importance for the neurite growth and guidance effects of HB-GAM, MK, and F-spondin. Based on the sequence similarities and probable structural similarities, we predict that heparin-type glycans are also important for biological roles of other TSR proteins, such as semaphorins F and G. Furthermore, thrombospondins interact with the surface of various cell types enhancing adhesion, spreading, and neurite outgrowth, migration, or proliferation (for review, see Ref. 49).

Cell surface glycosaminoglycans provide a potential source of a large array of different structures serving as binding partners in specific cell-cell or cell-matrix interactions (50). However, it is not known to what extent this structural diversity of glycosaminoglycans is used to regulate biological interactions. The occurrence of TSR domains in a large number of various proteins throughout the animal kingdom and their similarity in binding to heparin-type glycans suggest that this domain exploits the structural diversity of heparin epitopes in various biological interactions.

Acknowledgments—We acknowledge the SON NMR Large-Scale Facility in Utrecht (Dr. R. Boelens, Dr. M. Czisch, and Prof. Dr. R. Kaptein) for help and measuring time at high-field NMR instruments. We thank Dr. L. E. Kay for providing pulse sequences for heteronuclear spectroscopy. Prof. H. Benjamin Peng and Prof. Mart Saarma are acknowledged for helpful comments on the present study. The excellent technical assistance of Seija Lehto is gratefully acknowledged.

REFERENCES

1. Rauvala, H. (1989) *EMBO J.* **8,** 2933–2941
2. Merenmies, J., and Rauvala, H. (1990) *J. Biol. Chem.* **265,** 16721–16724
3. Matsubara, S., Tomomura, M., Kadomatsu, K., and Muramatsu, T. J. (1990) *J. Biol. Chem.* **265,** 9441–9443
4. Urios, P., Duprez, D., Le Caer, J.-P., Courtois, Y., Vigny, M., and Laurent, M. (1991) *Biochem. Biophys. Res. Commun.* **175,** 617–624
5. Rauvala, H., Vanhala, A., Castrén, E., Nolo, R., Raulo, E., Merenmies, J., and Panula, P. (1994) *Dev. Brain Res.* **79,** 157–176
6. Peng, H. B., Ali, A. A., Dai, Z., Daggett, D. F., Raulo, E., and Rauvala, H. (1995) *J. Neurosci.* **15,** 3027–3038
7. Mitsiadis, T. A., Salmivirta, M., Muramatsu, T., Muramatsu, H., Rauvala, H., Lehtonen, E., Jalkanen, M., and Thesleff, I. (1995) *Development* **121,** 37–51
8. Imai, S., Kaksonen, M., Raulo, E., Kinnunen, T., Fages, C., Meng, X., Lakso, M., and Rauvala, H. (1998) *J. Cell Biol.* **143,** 1112–1128
9. Raulo, E., Chernousov, M. A., Carey, D. J., Nolo, R., and Rauvala, H. (1994) *J. Biol. Chem.* **269,** 12999–13004
10. Nolo, R., Kaksonen, M., Raulo, E., and Rauvala, H. (1995) *Neurosci. Lett.* **191,** 39–42
11. Kinnunen, A., Kinnunen, T., Kaksonen, M., Nolo, R., Panula, P., and Rauvala, H. (1998) *Eur. J. Neurosci.* **10,** 635–648
12. Kinnunen, T., Raulo, E., Nolo, R., Maccarana, M., Lindahl, U., and Rauvala, H. (1996) *J. Biol. Chem.* **271,** 2243–2248
13. Kinnunen, T., Kaksonen, M., Saarinen, J., Kalkkinen, N., Peng, H. B., and Rauvala, H. (1998) *J. Biol. Chem.* **273,** 10702–10708
14. Rauvala, H., and Peng, H. B. (1997) *Prog. Neurobiol.* **52,** 127–144
15. Lauri, S. E., Kaukinen, S., Kinnunen, T., Ylinen, A., Imai, S., Kaila, K., Taira, T., and Rauvala, H. (1999) *J. Neurosci.* **19,** 1226–1235
16. Raulo, E., Julkunen, I., Merenmies, J., Pihlaskari, R., and Rauvala, H. (1992) *J. Biol. Chem.* **267,** 11408–11416
17. Hampton, B. S., Marshak, D. R., and Burgess, W. H. (1992) *Mol. Biol. Cell* **3,** 85–93
18. Li, Y. S., Milner, P. G., Chauhan, A. K., Watson, M. A., Hoffman, R. M., Kodner, C. M., Milbrandt, J., and Deuel, T. F. (1990) *Science* **250,** 1690–1694
19. Tezuka, K., Takeshita, S., Hakeda, Y., Kumegawa, M., Kikuno, R., and Hashimoto-Gotoh, T. (1990) *Biochem. Biophys. Res. Commun.* **173,** 246–251
20. Bohlen, P., and Kovesdi, I. (1991) *Prog. Growth Factor Res.* **3,** 143–157
21. Courty, J., Dauchel, M. C., Caruelle, D., Perderiset, M., and Barritault, D. (1991) *Biochem. Biophys. Res. Commun.* **180,** 145–151
22. Pervin, A., Gallo, C., Jandick, K. A., Han, X. J., and Linhardt, R. J. (1995) *Glycobiology* **5,** 83–95
23. Toida, T., Hileman, R. E., Smith, A. E., Vlahova, P. I., and Linhardt, R. J. (1996) *J. Biol. Chem.* **271,** 32040–32047
24. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* **114,** 6291–6293
25. Yamazaki T., Lee W., Arrowsmith C. H., Muhandiram D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116,** 11655–11666
26. Kay, L. E., Keifer, P., and Saarinen, T., (1992) *J. Am. Chem. Soc.* **114,** 10663–10665
27. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
28. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* **25,** 4876–4882
29. Muhandiram, D. R., and Kay, L. E., (1994) *J. Magn. Reson.* **103,** 203–216
30. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. A. (1992) *J. Magn. Reson.* **97,** 359–375
31. Wishart D. S., Sykes B. D., and Richards F. M. (1992) *Biochemistry* **31,** 1647–1651
32. Wishart D. S., and Sykes B. D. (1994) *J. Biomol. NMR* **4,** 171–180
33. Iwasaki, W., Nagata, K., Hatanaka, H., Inui, T., Kimura, T., Muramatsu, T., Yoshida, K., Tasumi, M., and Inagaki, F. (1997) *EMBO J.* **16,** 6936–6946
34. Hansen, M. R., Mueller, L., and Pardi, A. (1998) *Nat. Struct. Biol.* **5,** 1065–1074
35. Fath, M., VanderNoot, V., Kilpeläinen, I., Kinnunen, T., Rauvala, H., and Linhardt, R. J. (1999) *FEBS Lett.* **454,** 105–108
36. Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M. W., Hedgecock, E. M., and Culotti, J. G. (1992) *Cell* **71,** 289–299
37. Guo, N. H., Krutzsch, H. C., Negre, E., Vogel, T., Blake, D. A., and Roberts, D. D. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 3040–3044
38. Guo, N. H., Krutzsch, H. C., Negre, E., Zabrenetzky, V. S., and Roberts, D. D. (1992) *J. Biol. Chem.* **267,** 19349–19355
39. Gantt, S. M., Clavijo, P., Bai, X., Esko, J. D., and Sinnis, P. (1997) *J. Biol. Chem.* **272,** 19205–19213
40. Robson, K. J. H., Frevert, U., Reckmann, I., Cowan, G., Beier, J., Scragg, I. G., Takehara, K., Bishop, D. H. L., Pradel, G., Sinden, R., Saccheo, S., Müller H.-M., and Crisanti, A. (1995) *EMBO J.* **14,** 3883–3894
41. Naitza, S., Spano, F., Robson, K. J. H., and Crisanti, A. (1998) *Parasitol. Today* **14,** 479–484
42. Sawitzky, D. (1996) *Med. Microbiol. Immunol.* **184,** 155–161
43. Kuno, K., and Matsushima, K. (1998) *J. Biol. Chem.* **273,** 13912–13917
44. Cerami, C., Frevert, U., Sinnis, P., Takacs, B., Clavijo, P., Santos, M. J., and Nussenzweig, V. (1992) *Cell* **70,** 1021–1033
45. Sultan, A. A., Thathy, V., Frevert, U., Robson, K. J., Crisanti, A., Nussenzweig, V., Nussenzweig, R. S., and Menard, R. (1997) *Cell* **90,** 511–522
46. Muramatsu, H., and Muramatsu, T. (1991) *Biochem. Biophys. Res. Commun.* **177,** 652–658
47. Klar, A., Baldassare, M., and Jessell, T. M. (1992) *Cell* **69,** 95–110
48. Adams, R. H., Betz, H., and Püschel, A. W. (1996) *Mech. Dev.* **57,** 33–45
49. Bornstein, P. (1995) *J. Cell Biol.* **130,** 503–506
50. Lindahl, U., Kusche-Gullberg, M., and Kjellén, L. (1998) *J. Biol. Chem.* **273,** 24979–24982