ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS
Vol. 343, No. 1, July 1, pp. 92–100, 1997
Article No. BB970147

Pattern and Spacing of Basic Amino Acids in Heparin Binding Sites[1]

J. R. Fromm,* R. E. Hileman,* E. E. O. Caldwell,† J. M. Weiler,† and R. J. Linhardt*[,2]
*Division of Medicinal and Natural Products Chemistry Division and Department of Chemical and Biochemical Engineering, University of Iowa, Iowa City, Iowa 52242; and †Department of Internal Medicine, Iowa City VA Medical Center and University of Iowa, Iowa City, Iowa 52242

Received February 3, 1997, and in revised form April 3, 1997

Glycosaminoglycan (GAG)–protein interactions regulate a myriad of physiologic and pathologic processes, yet an understanding of how these molecules interact is lacking. The role of the pattern and spacing of basic amino acids (arginine (R) and lysine (K)) in heparin binding sites was investigated using peptide analogs as well as by examining known heparin binding sites. Peptides having the general structure R_nW (n = 3–9, where tyrosine (W) was added for peptide detection) were synthesized and their interaction with heparin was determined by isothermal titration calorimetry. Binding affinity increased with increasing number of R residues. A 9-mer of R (R_9W) bound as tightly to heparin as acidic fibroblast growth factor under physiologic conditions. Despite their high affinity for heparin, long stretches of basic amino acids are uncommon in heparin binding proteins. Known heparin binding sites most commonly contain single isolated basic amino acids separated by one nonbasic amino acid. Peptides having the structure, $H_3CCONH\text{-}GRRG_mRRG_{5\text{-}m}\text{-}CONH_2$ (denoted as the RRG_mRR peptide series) and $H_3CCONH\text{-}GRRRG_mRG_{5\text{-}m}\text{-}CONH_2$ (denoted as the $RRRG_mR$ peptide series), where m = 0–5, were synthesized to test the hypothesis that the spacing of basic amino acids in heparin binding sites is optimally arranged to interact with different GAGs. The peptides, in both the $\text{-}RRG_mRR\text{-}$ and $\text{-}RRRG_mR\text{-}$ peptide series, when m = 0, bound most tightly with heparin, as measured by affinity chromatography. In contrast, the $\text{-}RRG_mRR\text{-}$ peptide series interacted most tightly with heparan sulfate when m = 0 or 1, whereas the $\text{-}RRRG_mR\text{-}$ peptide series bound tightest when m = 3. These results are consistent with our understanding of heparin and heparan sulfate structure. A highly sulfated GAG, such as heparin, interacts most tightly with peptides (or peptide sequences within proteins) containing a complementary binding site of high positive charge density. Heparan sulfate, having fewer and more highly spaced negatively charged groups, interacts most tightly with a complementary site on a peptide (or peptide sequences with proteins) that has more widely spaced cationic residues. © 1997 Academic Press

The interaction of proteins with glycosaminoglycans (GAGs)[3] is of profound physiologic and pathologic importance. Diverse families of proteins such as growth factors, serine protease inhibitors, adhesion molecules, and lipoproteins interact with GAGs and the interactions often alter the activities of the proteins (1). Indeed, in some cases GAGs are required for protein activity (2–6). Although a number of proteins have been shown to bind heparin and other GAGs (7), how these species interact on a molecular level is not well understood. It is clear that positively charged, basic amino acids (arginine (R) and lysine (K)) play a critical role in the interaction of negatively charged sulfate and carboxylate groups of GAGs (1, 8, 9). An influential paper in 1989 by Cardin and Weintraub (10) suggested that positively charged amino acids tend to be arrayed on one face of an α-helix based on the observation that the pattern XBBBXXBX (where B is a basic and X a nonbasic amino acid) was found to be common in hepa-

[1] This work was supported in part by Grants HL52622 and GM38060 from the National Institutes of Health, a VA Merit Award, and fellowships from the Angiogenesis Foundation and the American Foundation for Pharmaceutical Education.

[2] Two whom correspondence and reprint requests should be addressed at University of Iowa PHAR S328, Iowa City, IA 52242. Fax: (319) 335-6634. E-mail: robert-linhardt@uiowa.edu.

[3] Abbreviations used: GAG, glycosaminoglycan; p-MBHA, p-methylbenzohydrylamine; t-Box, t-butyloxycarbonyl; FAB, fast atom bombardment; ITC, isothermal titration calorimetry; DTT, dithiotreitol; FGF, fibroblast growth factor; EC-SOD, extracellular superoxide dismutase; PDGF, platelet derived growth factor; HB-EGF, heparin binding endothelial growth factor.

0003-9861/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

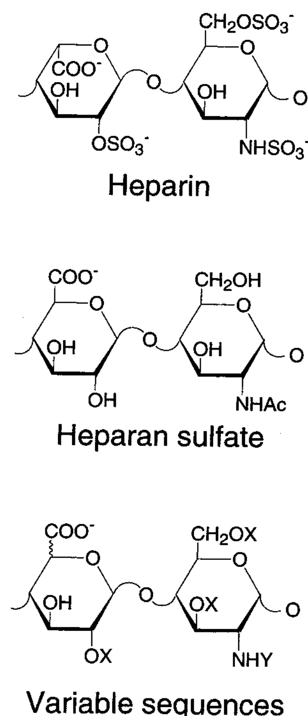

FIG. 1. Major repeating disaccharide units of heparin and heparan sulfate and variable disaccharide sequences (X = H or SO_3^- and Y = SO_3^-, CH_3CO or H) contained in both heparin and heparan sulfate.

rin binding sites; a second pattern, XBBXBX, aligns basic amino acids on one side of a β-strand. This paper, however, relied primarily on molecular modeling. Subsequently, Margalit and coworkers (11) used molecular modeling to suggest the importance of a crucial spacing of approximately 20 Å between positively charged amino acids in heparin binding sites, irrespective of whether the putative heparin binding element is in an α-helix or a β-sheet.

GAG heparin is derived from mast cells and basophils and is a highly sulfated linear polymer composed of polysaccharide chains of variable length and chemical structure (Fig. 1). Heparin's primary building blocks are iduronic acid and glucosamine and commercial chains contain an average of 2.7 sulfates per disaccharide residue (12). The positions and patterns of these sulfates may vary along the chains imparting structural variability. A related GAG, heparan sulfate, is found in the extracellular matrix and consists primarily of glucuronic acid and glucosamine; heparan sulfate contains less than 1 sulfate per disaccharide residue. Heparan sulfate has more structural variability than does heparin (13). Positively charged, basic amino acids in heparin and heparan sulfate binding proteins interact with the anionic groups in heparin and heparan sulfate, contributing a major component of the free energy (8, 14). Previous studies in our laboratory (8, 15) clearly demonstrated that R residues bind heparin with higher affinity than do K residues. The current study examined whether the ability of proteins to specifically interact with GAG sequences may be determined in part by the pattern of basic amino acids (in particular R). A combination of theoretical and experimental methods were used that suggested that the pattern as well as the spacing of basic amino acids were critical in determining affinity and specificity of GAG–protein interactions.

EXPERIMENTAL PROCEDURES

Materials

Heparan sulfate (12) and low molecular weight heparin (16), sodium salts, were purchased from Celsus Laboratories (Cincinnati, OH). The structural composition of these GAGs have been described in detail (12, 16). CNBr-Activated agarose and prepacked heparin-agarose columns were purchased from Sigma Chemical Co. (St. Louis, MO) Micro BCA assay kit, Reactigel, and *o*-phthaldehyde (OPA) solution were from Pierce (Rockford, IL). *t*-Butyloxycarbonyl (*t*-Boc) amino acids were from Advanced ChemTech (Louisville, KY). Resin, *p*-methylbenzhydrylamine (*p*-MBHA), was from Colorado Biotechnology Associates (Denver, CO). Trifluoroacetic acid was from Halocarbon Products of Augusta (Augusta, SC). Reversed phase C18 μBondapak columns were from Waters Chromatography (Milford, MA). All other reagents were from either Fisher Scientific (Pittsburgh, PA) or Aldrich Chemical (Milwaukee, WI).

Methods

Peptide synthesis and purification. Peptides were synthesized on *p*-MBHA resin using standard *t*-Boc solid phase chemistry (17, 18) or the t-bag technique using *p*-MBHA resin compartmentalized in polypropylene bags (19). The bags were pooled for the deblocking and base washing steps and were only separated for the coupling reactions. All amino acids were N-terminal blocked with *t*-Boc. Side chains were protected as: Arg (*N*-guanidino-toluenesulfonyl), Cys (S-4-methyl-benzyl); His (N^{im}-benzyloxymethyl); Lys (N-ϵ-2-chlorobenzyloxycarbonyl); Ser (*O*-benzyl); Thr (*O*-benzyl); and Tyr (*O*-2,6-dichlorobenzyl). Following the final deblocking step, the peptides were cleaved from the resin and their side chains were deprotected using a standard HF/anisole procedure (20). Up to 10 intact bags of resin were cleaved simultaneously in a compartmentalized reaction vessel from Multiple Peptide Systems (San Diego, CA). Residual anisole was removed with ethyl acetate before the peptides were extracted with 15% acetic acid. The resulting peptide preparations contained C-terminal amides. Crude peptides were analyzed by reverse phase HPLC (Waters μBondapak C18, 3.9 × 300 mm), starting at 1 ml/min with 0.1% trifluoroacetic acid in water and using linear gradients of 0 to 100%, 0.04% trifluoroacetic acid in acetonitrile. Preparative purification used and a Waters μBondapak C18 column (19 × 150 mm) and similar gradients.

Confirmation of peptide identity and purity. Fast atom bombardment (FAB) mass spectrometry was performed by the High Resolution Mass Spectrometry Facility of the Department of Chemistry at the University of Iowa. A ZAB HF VG analytical mass spectrometer was used to identify the molecular weight and to confirm the complete deprotection and peptide purity using either a glycerol or a thioglycerol matrix. Analysis of this mass spectral data also gave a partial sequence for each peptide that was consistent with its structure.

Preparation of heparan sulfate-agarose. CNBr-activated agarose (1 g) was swelled in 10 ml of 1 mM HCl for 20 min and then washed with 200 ml of 1 mM HCl. Coupling buffer (0.5 M sodium chloride, 0.1 M sodium bicarbonate, pH 9.5) was added to make a 50% slurry,

which was immediately poured into a heparan sulfate solution (6 ml coupling buffer containing 400 mg heparan sulfate). The gel was mixed for 2 h at 20°C, the supernatant was discarded, and the beads were washed with 4 M sodium chloride. The beads were blocked with 1 M ethanolamine, pH 9.0, at 25°C for 3.5 h and washed with 4 M sodium chloride. Toluidine blue assay (21) indicated the conjugate contained 0.75 mg heparan sulfate per milliliter of swelled matrix.

Heparin-agarose affinity chromatography of heterogeneous poly(R). Poly(R) (25 μg) was loaded onto a 2.5-ml heparin-agarose column (at 25°C) in 5 mM sodium phosphate, pH 7.4. The column was eluted by washing with 10 column volumes of loading buffer containing 50, 150, 300, 1000, and 2000 mM sodium chloride. Poly(R) in each sodium chloride wash was quantified by protein assay (22).

Heparin- and heparan sulfate-agarose affinity chromatography of synthetic peptides. Peptides were individually chromatographed on heparin-agarose (2.5 ml) or heparan sulfate-agarose (5 ml) columns. Briefly, columns were washed with 5 mM sodium phosphate buffer, pH 7.4, containing 1 M NaC1 and then 5 mM sodium phosphate, pH 7.4, buffer. Peptide (50 μg for heparin-agarose and 75 μg for heparan sulfate-agarose experiments) in 1 ml water was loaded, the column was washed with 10 column volumes of 5 mM sodium phosphate, pH 7.4, buffer and fractions (0.8 ml) were collected. Protein assay (22) confirmed that in all cases peptide bound under the equilibration conditions. The column was then eluted with a 0 to 1.0 M (for heparin-agarose) or 0 to 0.6 M (for heparan sulfate-agarose) sodium chloride gradient (12.5 column volumes) in 5 mM sodium phosphate, pH 7.4, buffer. Fractions (0.8 ml) were collected and analyzed for peptide by absorbance at 230 nm and protein assay (22). The salt concentration of the eluent was determined by measuring conductivity of the diluted (100-fold) fractions against a standard curve.

Isothermal titration calorimetry of peptides with low molecular weight heparin. Isothermal titration calorimetry (ITC) was used as previously described in detail (8). Briefly, 20 5-μl aliquots of low molecular weight heparin (0.5–1.0 mM) were injected every 200 s into a 1-ml cell containing peptide (50–100 μM) with stirring. Both interacting species were in either 10 mM sodium phosphate, pH 7.4, buffer containing 100 mM sodium chloride (-RRG$_m$RR- and -RRRG$_m$R- peptides) or 50 mM sodium phosphate, pH 7.4, buffer containing 100 mM sodium chloride (R$_n$W peptides). In control experiments, low molecular weight heparin was injected into the same buffer in the absence of peptide. Heats released on injection in both control and peptide containing experiments were integrated and the heats released from control experiments were subtracted to give corrected heats. These data were fitted to an equation that models the interaction, with the fitting of parameters n (number of ligand molecules binding to a macromolecule), ΔH (enthalpy change for reaction), and K_a (association constant) (23).

RESULTS

Heparin-agarose affinity fractionation poly(R). Initially we investigated the strength of polypeptide binding to heparin. Poly(R) homopolymer (containing chains of variable length with an average of 35 amino acids) was bound to heparin-agarose and eluted stepwise with salt washes (Fig. 2a). Over 20% of the poly(R) was retained on the column at a concentration of 1 M sodium chloride. This is comparable to the heparin binding affinity of many heparin binding proteins (24). Even after washing with 2 M sodium chloride a small amount of poly(R) still remained bound to the heparin-agarose column.

Isothermal titration calorimetry of R$_n$W. Arginine containing peptides (R$_n$W, n = 3–9) were synthesized

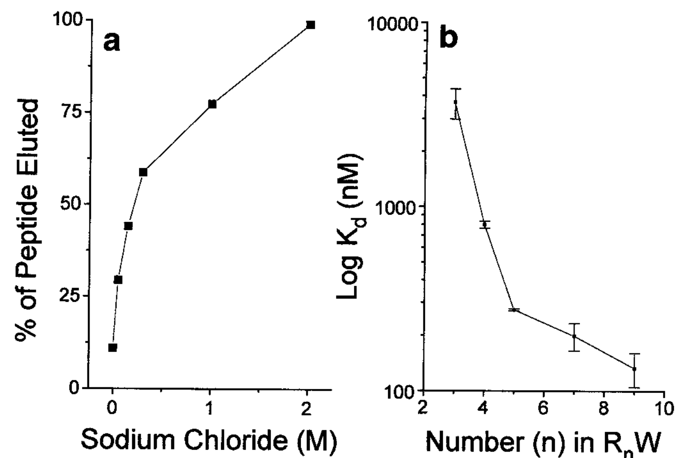

FIG. 2. Effect of number of arginine residues on peptide affinity to heparin. (a) Elution of poly(R) from heparin-agaroses as a function of sodium chloride concentration. (b) The dissociation constant, measured by ITC, of R$_n$W peptides for LMW heparin is plotted as a function of the number of R residues. Error bars are shown for standard deviation. The number of peptides bound per LMW heparin and ΔH for R$_3$W, R$_4$W, R$_5$W, R$_7$W and R$_9$W were: 6.1 ± 0.3, −35.2 ± 3; 4.2 ± 0.1, −39.1 ± 0.4; 3.5 ± 0.1, −37.8 ± 0.3; 2.8 ± 0.1, −34.8 ± 0.2; and 2.4 ± 0.1, −27.1 ± 0.6, respectively. Measurements were made in triplicate in 10 mM sodium phosphate, pH 7.4, buffer containing 100 mM sodium chloride at 25°C.

to study the interaction of defined arginine oligomers with heparin. Tryptophan (W) was included in each peptide to facilitate spectrophotometric quantification. Binding of peptides with low molecular weight heparin (MW$_{avg}$ 4,800) was studied using ITC (Fig. 2b). These data clearly demonstrate K_d is a function of the number of R residues. Interestingly, the K_d did not change appreciably once a cluster of 4 to 5 R residues is achieved. The R$_9$W peptide bound to heparin with affinity similar to heparin binding proteins (K_d of 133 ± 28 nM in 50 mM sodium phosphate containing 100 mM sodium chloride).

Investigation of patterns and spacing of basic amino acids in known heparin binding sites. Next, we examined patterns (clustering) of basic amino acids (i.e., -X$_m$B$_n$X$_m$-) in known heparin binding sites (Table I) to determine what patterns of spacing of basic amino are found in proteins. Sequences were selected based upon knowledge that they were known heparin binding sites as a result of: (a) ability of an analogous heparin binding site peptide to bind heparin, (b) inability of proteins mutated at the presumed binding site to bind heparin, (c) inability of proteins with chemically modified putative heparin binding site residues to bind heparin, and/or (d) ability of heparin to protect residues from chemical modification when the heparin binding protein is bound to heparin. Reported sequences containing more than 50 amino acids were not considered.

A single, isolated basic amino acid (i.e., -X$_m$BX$_m$-) was the most common pattern of basic amino acids in

TABLE I

Sequences of Selected Heparin Binding Sites

Protein[e]	Peptide sequences[f]	Criteria
bFGF (27)	^{93}FFFERLESNNYNTYRSRKYSSWYVALKR^{120}	*a*
Antistasin (28)	^{93}PNGLKRDKLGCEYCECRPKRKLIPRLS^{119}	*a*
Apo E (29)	^{144}LRKRLLRD^{151}	*a*
	^{211}GERLRARM^{218}	*a*
LPL (30, 31)	^{281}RKNRCNNLGYEINKVRAKR^{299}	*b*
EC-SOD (32)	^{205}REHSERKKRRRESECKAA^{222}	*b*
vWF (33)	^{565}YIGLKDRKRPSELRRIASQVKYA^{587}	*a*
NCAM (34)	^{131}TWKHKGRDVILKKDVRFI^{148}	*a*
Fibronectin (35–37)	^{6}RRARVTDATETTITISWRTKETETITGFQVDAIPANG^{41}	*b*
	^{1906}YEKPGSPPREVVPRPRPGV^{1924}	*a*
	^{1946}KNNQKSEPLIGRKKT^{1960}	*a*
Laminin (38, 39)	^{641}RYVVLPRPVCFEKGMNYTVR^{660}	*a*
	^{202}RIQNLLKITNLRIKFVK^{218}	*a*
	^{3010}KQNCLSSRASFRGCVRNLRLSR^{3032}	*a*
Vitronectin (40)	^{347}AKKQRFRHRNRKGYR^{361}	*a*
ATIII (41–43)	^{124}AKLNCRLYRKANKSSKLVSANR^{145}	*a, c, d*
PF4 (44, 45)	^{46}KDGRKICLDLQAPLYKKIIKKLLES^{70}	*a, c*
L-type C channel (8)	^{219}KGKMHKTCYY^{228}	*a*
	^{321}MGKMHKTCYN^{330}	*a*
aFGF (45–50)	^{115}KKHEAKNWFVGLKKGSCKRGP^{144}	*a, b, c*
Protein C inhibitor (51, 52)	^{264}SEKTLRKWLKMFKKRQLELY^{283}	*a, b*
90-kDa stress protein (53)	^{362}LYVR^{365}	*a*
	^{645}LRQK^{648}	*a*
Thrombospondin (54–56)	^{23}RKGSGRRLVK^{30}	*b*
	^{77}RQMKKTR^{83}	*b*
TGF β1 (57)	^{23}DFRKDLGWKWIHEPKGYHA^{41}	*a*
Apo B100 (58)	^{3144}LSVKAQYKKNKHRHSI^{3159}	*a*
	^{3352}YKLEGTTRLTRKRGLKLATA^{3371}	*a*
PDGF-A (59, 60)	^{194}GRPRESGKKRKRKRLKPT^{211}	*a*
Xanthine oxidase (61)	^{781}LGVPANRIVVRVKRM^{795}	*a*
	^{1106}KKKNPSGSWEDWVTAAY^{1122}	*a*
Glia derived nexin (62)	^{63}RYNVNGVGKVLKKINKAIVSKKNK^{86}	*b*
TFPI (63)	^{212}GKCRPFKYSGCGGNENNFTSKQECLRACKKGF^{243}	*a, b*
AAMP (64)	^{14}RRLRRMESESES^{25}	*a*
IGFBP-5 (65)	^{221}RKGFYKRKQCKPSRGRKR^{238}	*a*
IGFBP-3 (66)	^{215}KKGFYKKKQCRPSKGRKR^{232}	*a*
HB-EGF (67)	^{21}KRKKKGKGLGKKRDPCLRKYK^{41}	*a, b*

[a] Ability of an analogous heparin binding peptide to bind heparin.
[b] Inability of proteins mutated at the presumed binding site to bind heparin.
[c] Inability of chemically modified putative heparin binding residues to bind heparin.
[d] Ability of heparin to protect residues from chemical modification when the heparin binding protein is bound to heparin.
[e] References for the sequences in parens.
[f] Superscripts refer to amino acids positions in the full-length protein.

protein heparin binding sites, followed by clusters of 2 (i.e., -X_mBBX_m-) or 3 (i.e., -X_mBBBX_m-) basic amino acids (Fig. 3a). Patterns with clusters of 5, 6, or 7 basic amino acids were uncommon. Sequences with clusters of 4 basic amino acids were not found in this sample of heparin binding sites. Despite the fact that extended clusters of basic amino acids can afford extremely high affinity for heparin binding (see Fig. 2) such extended clusters appear to be uncommon in the heparin binding sites of proteins.

The spacing of nonbasic amino acids (i.e., -$B_nX_mB_n$-) between basic amino acid clusters in heparin binding sites was also examined (Fig. 3b). The most common spacing was a single nonbasic amino acid between two basic amino acid clusters (i.e., -B_nXB_n-). A spacing of nonbasic 2 amino acids between the basic clusters (i.e., -$B_nX_2B_n$-) was the second most commonly observed pattern. Although spacings of 3, 4, and 5 nonbasic amino acids are still relatively common, spacings of greater than 7 nonbasic amino acids are uncommon.

Heparin-agarose affinity fractionation peptides with defined patterns and spacing of basic amino acids. Peptides with the general sequence H_3CCONH-GRRG$_m$RRG$_{5\text{-}m}$-$CONH_2$ (denoted -RRG$_m$RR- peptide series) and H_3CCONH-GRRRG$_m$RG$_{5\text{-}m}$-$CONH_2$ (denoted -RRRG$_m$R- peptide series), where $m = 0–5$, were synthesized to assess the importance of spacing between basic residue clusters. Arginine was again chosen as the basic amino acid because

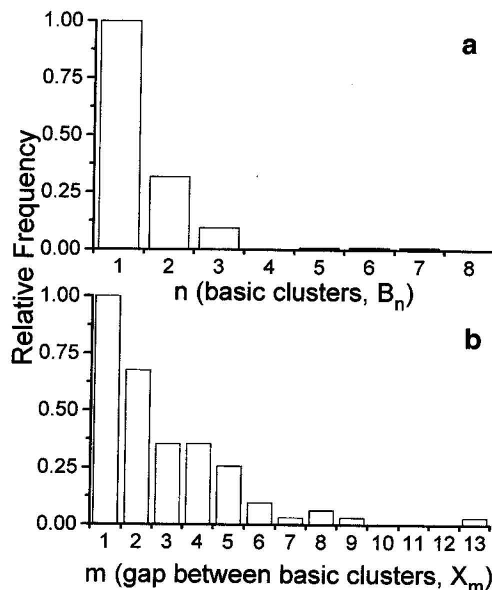

FIG. 3. Patterns and spacing of basic amino acids in heparin binding sites. (a) Relative frequency of different basic amino acid (B) clusters in heparin binding sites. B_{1-8} represent 1–8 basic amino acids in a row. (b) Relative frequency of nonbasic amino acids (X) between clusters of basic amino acids in heparin binding sites. X_{1-13} represents 1–13 nonbasic amino acids spacing between basic clusters.

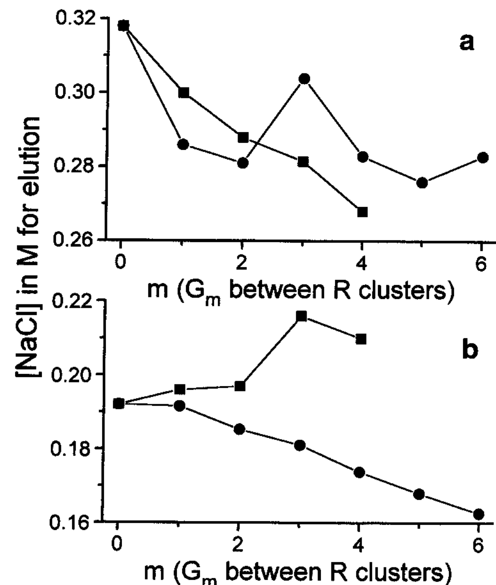

FIG. 4. Affinity fractionation of peptides with defined patterns and spacing of basic amino acids. (a) Affinity for heparin-agarose (measured by concentration of NaC1 required for elution) is plotted as a function of m glycines (G_m) between arginine (R) clusters for the RRG_mRR peptide series (●) and $RRRG_mR$ peptide series (■). (b) Affinity for heparan sulfate-agarose (as measured by concentration of NaC1 required for peptide elution) is plotted as a function of m glycines (G_m) between arginine (R) clusters for the RRG_mRR peptide series (●) and $RRRG_mR$ peptide series (■). The data shown is represenative of two separate determinations for each peptide series.

it results in tighter interaction with GAGs than does lysine (8, 9). Glycine (G) was selected as the nonbasic amino acid spacer residue as it is the simplest nonbasic amino acid and does not contain a side chain that might interact with GAGs. The sequences of these peptides are shown in Table II. Peptides contained four basic residues as preliminary studies had demonstrated that four R residues were sufficient to achieve strong binding to heparin (Fig. 2b). All of the peptides had identical amino acid compositions and thus, identical charges of +4. Both amino and carboxy termini of the peptides were blocked as amides to resemble more closely sequences found in heparin-binding proteins.

Heparin-agarose affinity chromatography was used to assess the affinity of the peptides for heparin. The

TABLE II

Sequences of Peptides Used to Assess Importance of Spacing in Heparin Binding Sites

RRG_mRR series	$RRRG_mR$ series	*m*
Ac-GRRRRGGGGG-Amide	Ac-GRRRRGGGGG-Amide	0
Ac-GRRGRRGGGG-Amide	Ac-GRRRGRGGGG-Amide	1
Ac-GRRGGRRGGG-Amide	Ac-GRRRGGRGGG-Amide	2
Ac-GRRGGGRRGG-Amide	Ac-GRRRGGGRGG-Amide	3
Ac-GRRGGGGRRG-Amide	Ac-GRRRGGGGRG-Amide	4
Ac-GRRGGGGGRR-Amide		5
Ac-RRGGGGGGRR-Amide		6

salt concentration required to elute the peptides represents a measure of this affinity (25). The affinity of the peptides, as a function of the number of glycine residues between the basic clusters, is shown in Fig. 4a. For the -RRG_mRR- pattern, the peptides where $m = 0$ and $m = 3$, had the highest affinity for heparin-agarose. A different pattern was seen for the -$RRRG_mR$- series. The peptide without a glycine spacer (i.e., H_3CCONH-GRRRRGGGGG-$CONH_2$) bound most tightly to heparin; affinity decreased as the number of glycine spacers increased.

Heparan sulfate-agarose affinity fractionation of peptides with defined patterns and spacing of basic amino acids. Peptides -RRG_mRR- and -$RRRG_mR$- were next examined for affinity to heparan sulfate-agarose as measured by salt concentration required for elution (Fig. 4b). Both peptide series bound with less affinity to heparan sulfate-agarose than to heparin-agarose. Interestingly, the pattern of elution for heparan sulfate-agarose differed significantly from the pattern seen for heparin-agarose affinity experiments (Fig. 4b). In the -RRG_mRR- series, $m = 0$ and 1 (i.e., -RRRR- and -RRGRR-) had the highest affinity with affinity decreased as m increased from 2 to 6. With the -$RRRG_mR$- peptide series, an entirely different affinity pattern was observed. Peptides with $m = 0–2$ had similar affinities, whereas, when $m = 3$ and 4, peptides bound heparan

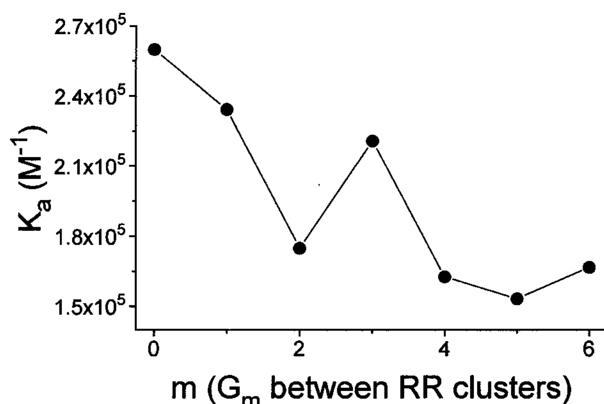

FIG. 5. Microtitration calorimetry measurement on the binding of the RRG_mRR peptide series to low molecular weight heparin. The association constant (K_a) is plotted as a function m of glycines (G_m) between the RR clusters. Data points are averaged from three separate determinations.

sulfate with considerably higher affinity. The pattern and spacing of basic amino acids that recognized heparan sulfate most avidly were -RRRGGGR- and -RRRGGGGR-.

Isothermal titration calorimetry to measure binding of -RRG_mRR- peptides to heparin. ITC was employed as a second, independent method to confirm the pattern of the -RRG_mRR- peptides series observed in the heparin-agarose affinity chromatography experiment as well as to examine the thermodynamics of the peptide-heparin interaction. This technique affords values for K_a, ΔH (enthalpy or heat of interaction), and the number of interactions between species in a single experiment. Thermodynamic data (Table III) were used to obtain association constants that are plotted in Fig. 5 as a function of the number of glycine spacers (m) between the basic R clusters (-RRG_mRR-). The shape of the curve is remarkably similar to that observed in the affinity chromatography experiments with the -RRG_mRR- peptides series binding to heparin-agarose (Fig. 4a). Again the peptide with no G spacer (-RRRR-) bound heparin tightest with the peptide containing the -RRGGGRR- sequence also showing high affinity for heparin.

DISCUSSION

It is becomingly increasingly clear that GAGs alter the activities of many heparin binding proteins (1). GAGs appear to be required for some proteins to mediate their biologic effects, as observed for the fibroblast growth factor (FGF) family (2–4). The interaction of FGF with heparin is well studied and a cocrystal structure of a heparin tetrasaccharide and a hexasaccharide binding to basic FGF has recently been published (26). Less effort has been made to understand these heparin protein interactions in a more general sense. One study suggested that defined patterns of amino acids, XBBBXXBX and XBBXBX (where X is a hydropathic amino acid and B is a basic amino acid) are common elements in GAG binding proteins (10). The XBBBXXBX and XBBXBX patterns recognize GAGs if the sequence lies in an α-helix or β-strand, respectively, thus aligning the basic amino acids linearly on one face of the α-helix or β-strand. A second study suggested that a 20 Å spacing between basic amino acids on opposite faces of an α-helix or β-strand was critical in GAG-protein recognition (11).

Experiments in our laboratories have demonstrated that R was the most common amino acid in randomly synthesized peptides that bound with high affinity to heparin (15). In contrast, when affinity for heparan sulfate was assessed, R was only the fourth most common amino acid with significantly more G and serine (S) residues observed in the same randomly synthesized peptide library (15). This enigmatic result prompted the current study to understand the role of spacing of amino acids in heparin binding sites.

This study examined known binding sites in proteins as well as the interaction of model peptides having variable patterns and spacing of basic amino acids with GAGs to help understand GAG-protein interactions. Preliminary studies showed that an increase in R residues in R_nW lead to increased binding affinity, and that 4 to 5 R residues afforded nearly as tight binding as 9 R residues. R_9W was shown to bind with very high affinity to heparin (K_d of 133 ± 28 nM), comparable to bFGF binding to heparin. Indeed, large poly(R) oligomers bound so tightly to heparin-agarose that they could not be eluted completely off the columns even with a 2 M sodium chloride solution. Known heparin binding sites were next examined to determine if sites analogous to R_9W were found in heparin binding proteins. Surprisingly, clusters of 1, 2, or 3 basic amino acids were the most common clusterings observed in known heparin binding sites. Spacing of such clusters

TABLE III

Thermodynamic Data for Binding of RRG_mRR Peptide Series to Heparin[a]

m	K_d (μM)	K_a (M^{-1})	No. of peptides per LMW heparin	ΔH (kcal/mol)
0	3.85	259740	9.49	−10.05
1	4.27	234192	8.72	−10.80
2	5.72	174825	11.40	−10.88
3	4.53	220751	9.90	−12.00
4	6.15	162602	7.70	−12.20
5	6.53	153139	7.80	−12.30
6	6.00	166667	11.90	−14.00

[a] The data are averaged from three separate determinations. Measurements were made at 25°C in 10 mM sodium phosphate, 100 mM NaCl, pH 7.4.

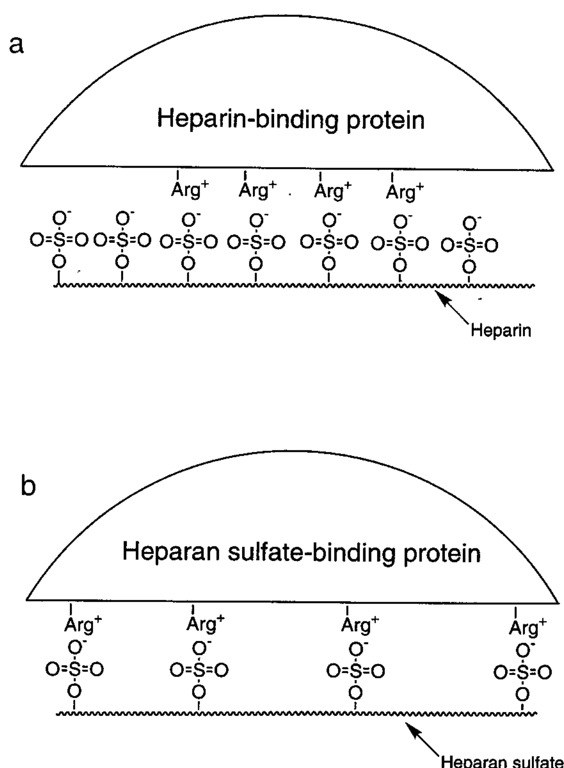

FIG. 6. Schematic diagram showing the interaction of a protein with (a) heparin having a high charge density and (b) heparan sulfate having a lower charge density. While both diagrams show the same number of ion-pairing interactions the spacing of these interactions is different leading to different specificities.

up to 5 residues apart was relatively common, although close (1 or 2 nonbasic amino acids spacing between basic clusters) arrangement of these clusters was most commonly observed. While GAG binding sites made of extended clusters (-$XR_{>4}X$-) of basic amino acids do occur, they are relatively uncommon and appear not to represent the mechanism chosen in nature to achieve high affinity binding.

Peptides with shorter basic clusters separated by G residues were next examined. The peptide with the highest heparin affinity in the RRG_mRR peptide series was that with $m = 0$ (i.e., a -RRRR- sequence). The second highest affinity was observed for the peptide with $m = 3$ (-RRGGGRR-) as determined by affinity chromatography (Fig. 4a). This peptide also showed high affinity toward LMW heparin using ITC (Fig. 5). Heparin is a highly sulfated polysaccharide with an average of 2.7 sulfates per disaccharide unit (12). For a protein or peptide to ion-pair tightly with heparin, a high positive charge density (a cluster of four R residues) might provide an ideal complementary binding site (Fig. 6a). In the -$RRRG_mR$- peptide series, heparin affinity decreased with increased spacing consistent with its complimentary pairing to the highly sulfated heparin polymer. It is less clear why a peptide with $m = 3$ (-RRGGGRR-) in the RRG_mRR series had enhanced heparin affinity. Heparin is known to form a helix in solution (24). The peptide with $m = 3$ may have been the proper size to bind one face and span one turn of the heparin helix. Thus, a pattern of -BBXXXBB-, where B is a basic amino acid and X a nonbasic amino acid, might also specifically recognize highly sulfated heparin polymer.

The binding of the -RRG_mRR- peptide series to heparan sulfate shows that peptides with $m = 0$ and 1 bind with similar affinity. As spacing between the clusters increases the affinity falls off. In known heparin binding sites, the most common spacing between basic amino acid clusters is one nonbasic amino acid (Fig. 3b). Perhaps this pattern is commonly observed in heparin-binding proteins (Fig. 3) because it is optimally spaced for heparan sulfate recognition. It should be reemphasized that most heparin-binding proteins are heparan sulfate-binding proteins, as heparan sulfate is found on the cell surface and in the extracellular matrix and is most certainly the endogenous ligand (7). The -$RRRG_mR$- peptide series interacted with heparan sulfate with a markedly different pattern of affinity. The binding affinity increased as spacing between the basic clusters was increased. Heparan sulfate has less than 1 sulfate per disaccharide, and thus has a much lower charge density than heparin (12). The greater spacing of sulfate groups in heparan sulfate may make it pair better with the more widely spaced basic clusters in this series. Heparan sulfate has regions of high and low sulfation. This variation in sulfation pattern might accommodate a wider range of spacing motifs than the structurally less complex heparin molecule.

Our model for the role of spacing in protein-GAG sequence recognition is shown in Figs. 6a and 6b). The Arg^+ (or Lys^+) cations of the protein are tightly clustered in the real heparin-binding proteins (such as those found in the mast cell granule having access to intracellular heparin) to ion-pair optimally with the complementary tightly clustered sulfate groups in the polymer (Fig. 6a). Heparan sulfate has a lower anionic charge density and would be expected to interact tightly with heparan sulfate binding proteins (such as those found in extracellular matrix where heparan sulfate is located) with sites containing positive charge groups that are "spread out" (Fig. 6b).

Heparin-binding proteins containing four or more basic amino acids in a row are uncommon. However, extracellular superoxide dismutase (EC-SOD), platelet derived growth factor (PDGF-A), heparin binding endothelial growth factor (HB-EGF), and possibly the proteins found in mast cell granules, all contain binding sites with greater than four basic amino acids in a row (Table I) and thus are mimicked by the peptide with -RRRR- cluster in its sequence. The present study suggests that these sites would interact most

effectively with heparin or highly sulfated regions in heparan sulfate.

Many heparin-binding sites (Table I) contain two or more sets of spaced basic amino acid clusters (i.e., -BB-X_m-BB-X_m-BB-X_m-BB-). Multiple clusters of basic amino acids in a protein that has the proper spacing should result in much stronger binding. We have previously shown that R residues in heparin binding sites bind more tightly to sulfate groups than do K residues and suggested that binding site affinity might be tailored by the "correct" ratio of R to K for optimum binding affinity and physiologic response. Similarly, the pattern and spacing of basic amino acids may represent a second means by which affinity of heparin-binding sites is determined.

REFERENCES

1. Jackson, R. L., Busch, S. J., and Cardin, A. D. (1991) *Physiol. Rev.* **71,** 481–539.
2. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., and Ornitz, D. M. (1991) *Cell* **64,** 841–848.
3. Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E., and Leder, P. (1992) *Mol. Cell. Biol.* **12,** 240–247.
4. Spivak-Kroizman, T., Lemmon, M. A., Dikic, I., Ladbury, J., Pinchasi, D., Huang, J., Jaye, M., Crumley, G., Schlessinger, J., and Lax, I. (1994) *Cell* **79,** 1015–1024.
5. Sudhalter, J., Folkman, J., Svahn, C. M., Bergenda, K., and D'Amore, P. A. (1989) *J. Biol. Chem.* **264,** 6892–6897.
6. Mueller, S. N., Thomas, K. A., Salvo, J. D., and Levine, E. M. (1989) *J. Cell. Physiol.* **140,** 439–448.
7. Linhardt, R. J. (1997) *in* Carbohydrates as Drugs (Witczak, Z. B., and Nieforth, K. A., Eds.), pp. 277–341, Marcel Dekker, New York.
8. Fromm, J. R., Hileman, R. E., Caldwell, E. E. O., Weiler, J. M., and Linhardt, R. J. (1995) *Arch. Biochem. Biophys.* **323,** 279–287.
9. Mascotti, D. P., and Lohman, T. M. (1995) *Biochemistry* **34,** 2908–2915.
10. Cardin, A. D., and Weintraub, H. J. R. (1989) *Arteriosclerosis* **9,** 21–32.
11. Margalit, H., Fischer, N., and Ben-Sasson, S. A. (1993) *J. Biol. Chem.* **268,** 19228–19231.
12. Griffin, C. C., Linhardt, R. J., Van Gorp, C. L., Toida, T., Hileman, R. E., Schubert, R. L., and Brown, S. E. (1995) *Carbohydr. Res.* **276,** 183–197.
13. Gallagher, J. T., and Walker, A. (1985) *Biochem. J.* **230,** 665–674.
14. Lohman, T. M., and Mascotti, D. P. (1992) *Methods Enzymol.* **212,** 400–424.
15. Caldwell, E. E. O., Nadkarni, V. D., Fromm, J. R., Linhardt, R. J., and Weiler, J. M. (1996) *Int. J. Biochem. Cell Biol.* **28,** 203–216.
16. Linhardt, R. J., Loganathan, D., Al-Hakim, A., Wang, H.-M., Walenga, J. M., Hoppensteadt, D., and Fareed, J. (1990) *J. Med. Chem.* **33,** 1639–1645.
17. Merrifield, R. B. (1963) *J. Am. Chem. Soc.* **85,** 2149–2154.
18. Houghten, R. A., Ostresh, J. M., and Klipstein, F. A. (1984) *Eur. J. Biochem.* **145,** 157–162.
19. Houghten, R. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 5131–5135.
20. Houghten, R. A., Bray, M. K., Degraw, S. T., and Kirby, C. J. (1986) *Int. J. Pept. Protein Res.* **27,** 673–678.
21. Smith, P. K., Mallia, A. K., and Hermanson, G. T. (1980) *Anal. Biochem.* **109,** 466–473.
22. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) *Anal. Biochem.* **150,** 76–85.
23. Freire, E., Mayorga, O. L., and Straume, M. (1990) *Anal. Chem.* **62,** 950A–959A.
24. Farouqui, A. A., and Horrocks, L. A. (1983) *Adv. Chromatogr.* **23,** 127–148.
25. Thompson, L. D., Pantoliano, M. W., and Springer, B. A. (1994) *Biochemistry* **33,** 3831–3840.
26. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J., and Rees, D. C. (1996) *Science* **271,** 1116–1120.
27. Baird, A., Schubert, D., Ling, N., and Guillemin, R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2324–2328.
28. Manley, G. D., Owen, T. J., Krstenanshy, J. L., Brankamp, R. G., and Cardin, A. D. (1992) Heparin and Related Polysaccharides (Lane, D. A., Ed.), Plenum Press, New York.
29. Cardin, A. D., Hirose, N., Blankenship, D. T., Jackson, R. L., and Harmony, J. A. K. (1986) *Biochem. Biophys. Res. Commun.* **134,** 783–789.
30. Berryman, D. E., and Bensadoun, A. (1993) *J. Biol. Chem.* **268,** 3272–3276.
31. Hata, A., Ridinger, D. N., Sutherland, S., Emi, M., Shuhua, Z., Myers, R. L., Ren, K., Cheng, T., Inoue, I., Wilson, D. E., Iverius, P.-H., and Lalouel, J.-M. (1993) *J. Biol. Chem.* **268,** 8447–8457.
32. Sandström, J., Carlsson, L., Marklund, S. L., and Edlund, T. (1992) *J. Biol. Chem.* **267,** 18205–18209.
33. Sobel, M., Soler, D. F., Kermode, J. C., and Harris, R. B. (1992) *J. Biol. Chem.* **267,** 8857–8862.
34. Kallapur, S. G., and Akeson, R. A. (1992) *J. Neurosci. Res.* **33,** 538–548.
35. Barkalow, F. J. B., and Schwarzbauer, J. E. (1991) *J. Biol. Chem.* **266,** 7812–7818.
36. Haugen, P. K., McCarthy, J. B., Skubitz, A. P. N., Furcht, L. T., and Letourneau, P. C. (1990) *J. Cell Biol.* **111,** 2733–2745.
37. McCarthy, J. B., Skubitz, A. P. N., Zhao, Q., Yi, X., Mickelson, D. J., Klein, D. J., and Furcht, L. T. (1990) *J. Cell Biol.* **110,** 777–787.
38. Charonis, A. S., Skubitz, A. P. N., Koliakos, G. G., Reger, L. A., Dege, J., Vogel, A. M., Wohlhueter, R., and Furcht, L. T. (1988) *J. Cell Biol.* **107,** 1253–1260.
39. Kouzi-Koliakos, K., Koliakos, G. G., Tsilibary, E. C., Furcht, L. T., and Charonis, A. S. (1989) *J. Biol. Chem.* **264,** 17971–17978.
40. de Boer, H. C., Preissner, K. T., Bouma, B. N., and de Groot, P. G. (1992) *J. Biol. Chem.* **267,** 2264–2268.
41. Lellouch, A. C., and Lansbury, P. T. J. (1992) *Biochemistry* **31,** 2279–2285.
42. Smith, J. W., and Knauer, D. J. (1987) *J. Biol. Chem.* **262,** 11964–11972.
43. Smith, J. W., Dey, N., and Knauer, D. J. (1990) *Biochemistry* **29,** 8950–8957.
44. Loscalzo, J., Melnick, B., and Handin, R. I. (1985) *Arch. Biochem. Biophys.* **240,** 446–455.
45. Rucinski, B., Niewiarowski, S., Strzyzewski, M., Holt, J. C., and Mayo, K. H. (1990) *Thromb. Haemost.* **63,** 493–398.
46. Wong, P., Hampton, B., Szylobryt, E., Gallagher, A. M., Jaye, M., and Burgess, W. H. (1995) *J. Biol. Chem.* **270,** 25805–25811.
47. Harper, J. W., and Lobb, R. R. (1988) *Biochemistry* **27,** 671–678.

48. Jaye, M., Howk, R., Burgess, W., Ricca, G. A., Chiu, I. M., Ravera, M. W., O'Brien, S. J., Modi, W. S., Maciag, T., and Drohan, W. N. (1986) *Science* **233,** 541–545.

49. Burgess, W. H., Shaheen, A. M., Ravera, M., Jaye, M., Donohue, P. J., and Winkles, J. A. (1990) *J. Cell. Biol.* **111,** 2129–2138.

50. Burgess, W. H., Shaheen, A. M., Hampton, B., Donohue, P. J., and Winkles, J. A. (1991) *J. Cell. Biochem.* **45,** 131–138.

51. Shirk, R. A., Elisen, M. G. L. M., Meijers, J. C. M., and Church, F. C. (1994) *J. Biol. Chem.* **269,** 28690–28695.

52. Pratt, C. W., and Church, F. C. (1992) *J. Biol. Chem.* **267,** 8789–8794.

53. Itoh, H., and Tashima, Y. (1993) *Int. J. Biochem.* **25,** 157–161.

54. Lawler, J., Duquette, M., and Ferro, P. (1991) *J. Biol. Chem.* **266,** 8039–8043.

55. Dixit, V. M., Hennessy, S. W., Grant, G. A., Rotwein, P., and Frazier, W. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5449–5453.

56. Murphy-Ulrich, J. E., Gurusiddappa, S., Frazier, W. A., and Hook, M. (1993) *J. Biol. Chem.* **268,** 26784–26789.

57. McCaffrey, T. A., Falcone, D. J., and Du, B. (1992) *J. Cell. Physiol.* **152,** 430–440.

58. Weisgraber, K. H., and Rall, S. C. J. (1987) *J. Biol. Chem.* **262,** 11097–11103.

59. Khachigian, L. M., Owensby, D. A., and Chesterman, C. N. (1992) *J. Biol. Chem.* **267,** 1660–1666.

60. Fager, G., Camejo, G., Olsson, U., Ostergren-Lunden, G., and Bondjers, G. (1992) *In Vitro Cell. Dev. Biol.* **28A,** 176–180.

61. Fukushima, T., Adachi, T., and Hirano, K. (1995) *Biol. Pharm. Bull.* **18,** 156–158.

62. Stone, S. R., Brown-Luedi, M. L., Rovelli, G., Guidolin, A., McGlynn, E., and Monard, D. (1994) *Biochemistry* **33,** 7731–7735.

63. Enjyoji, K., Miyate, T., Kamikubo, Y., and Kato, H. (1995) *Biochemistry* **34,** 5725–5735.

64. Beckner, M. E., Krutzsch, H. C., Stracke, M. L., Williams, S. T., Gallardo, J. A., and Liotta, L. A. (1995) *Cancer Res.* **55,** 2140–2149.

65. Arai, T., Parker, A., Busby, W. J., and Clemmons, D. R. (1994) *J. Biol. Chem.* **269,** 20388–20393.

66. Booth, B. A., Boes, M., Andress, D. L., Dake, B. L., Kiefer, M. C., Maack, C., Linhardt, R. J., Bar, K., Caldwell, E., Weiler, J., and Bar, R. S. (1995) *Growth Regul.* **5,** 1–17.

67. Thompson, S. A., Higashiyama, S., Wood, K., Pollitt, N. S., Damm, D., McEnroe, G., Garrick, B., Ashton, N., Lau, K., Hannock, N., Klagsbrun, M., and Abraham, J. A. (1994) *J. Biol. Chem.* **269,** 2541–2549.