Pacific Symposium on Biocomputing 5:152-164 (2000)

# Pinpointing the putative heparin/sialic acid-binding residues in the 'sushi' domain 7 of factor H: a molecular modeling study.

S. RANGANATHAN
*Australian Genomic Information Centre, C80 ATP,*
*University of Sydney, Sydney NSW 2006, Australia*

D.A. MALE, R.J. ORMSBY, E. GIANNAKIS, D.L. GORDON
*Department of Microbiology and Infectious Diseases,*
*Flinders University of South Australia, Bedford Park SA 5042, Australia*

Factor H, a secretory glycoprotein comprising 20 short consensus repeat (SCR) or 'sushi' domains of about 60 amino acids each, is a regulator of the complement system. The complement-regulatory functions of factor H are targeted by its binding to polyanions such as heparin/sialic acid, involving SCRs 7 and 20. Recently, the SCR 7 heparin-binding site was shown to be co-localized with the *Streptococcus* Group A M protein binding site on factor H (T.K. Blackmore *et al., Infect. Immun.* **66**, 1427 (1998)). Using sequence analysis of all heparin-binding domains of factor H and its closest homologues, molecular modeling of SCRs 6 and 7, and surface electrostatic potential studies, the residues implicated in heparin/sialic acid binding to SCR 7 have been localized to four regions of sequence space containing stretches of basic as well as histidine residues. The heparin-binding site is spatially compact and lies near the interface between SCRs 6 and 7, with residues in the interdomain linker playing a significant role.

## 1 Introduction

Complement factor H (fH) is an important member of the regulators of complement (C) activation (RCA) family of proteins, encoded by a cluster of genes on chromosome 1.[1] C-regulation by fH is effected by inhibition of the formation and acceleration of the decay of the alternate pathway C3 convertase (C3bBb),[2,3] and by serving as a cofactor for the C3b-cleaving enzyme complement factor I.[4] fH also shows chemotactic activity for monocytes[5] and may interact with the extracellular matrix and leucocytes.[1,6] Genetic deficiency of fH has been implicated in diseases such as glomerulonephritis[7] and hemolytic uraemic syndrome.[8] One of the critical roles ascribed to fH is to distinguish self from non-self by selectively allowing continual C activation on foreign particles and not on host cells. C activation is inhibited on host cells to which fH is bound,[9] which in turn is determined by the amount of membrane-associated sialic acid present on cell surfaces.[10] fH binds more strongly to heparin (which contains several sialic acid residues) than to sialic acid itself, so that heparin-binding to fH has been used as an investigative tool in several experimental studies.[10]

fH is a prototype of proteins with a modular structure,[11] containing a tandem array of homologous units called short consensus repeat (SCR or 'sushi') regions,

which have now been recognized in 12 complement proteins and many non-complement proteins including blood clotting factor XIIIb, the interleukin-2 receptor α-chain, and cell adhesion molecules such as endothelial leukocyte adhesion molecule-1 and leukocyte adhesion molecule-2.[12] Each SCR is characterized[11] by conserved tyrosine, proline and glycine residues and by the presence of four conserved cysteine residues, forming two disulfide bridges with 1-3 and 2-4 connectivity.[13] The 20 SCR domains of human fH (hfH) are connected by short linkers of three to eight residues, arranged in a head-to-tail fashion, resembling a string of beads.[14]

Identification of the heparin-binding site on hfH has been the subject of several experimental investigations, to test the hypothesis that the heparin/sialic acid-binding capacity of fH is essential to self/non-self recognition in the alternative pathway. Of the 20 SCRs in hfH, SCRs 7 and 20 alone exhibit heparin-binding activity,[15,16] with the *Streptococcus* Group A M protein sharing the heparin-binding site on SCR 7.[17] The binding of fH to M protein prevents complement activation and protects the pathogen from immune responses.[18] In addition to understanding how the heparin binding function of fH mediates self/non-self identification and the virulence of microbial pathogens, fH has therapeutic potential for reducing C activation by hemodialysis and cardiac bypass circuits. Heparin coating has been used to decrease C activation by bioincompatible surfaces but it is unclear if these effects are mediated by fH.[19,20]

While SCRs 7 and 20 of hfH contain heparin-binding sites, their functionality is dependent on the presence of N-terminal flanking SCRs.[16] SCR 7 requires at least one preceding SCR (by analogy to the heparin-binding SCR 2 of human fH-related protein fHR-3[16] containing domains homologous to hfH SCRs 6,7,9,19,20) while SCR 20 requires at least the 2 preceding SCRs 18 and 19. In an effort to study interdomain interaction and contact regions, the NMR structures of hfH SCRs 15 and 16 (fH1516)[21] and domains 3 and 4 from the Vaccinia virus complement control protein (vcp34)[22] were determined. The structures obtained[21,22] show that while each SCR folds essentially into a globular domain, there is considerable variation in their relative orientation in a two-domain protein. From these structural studies, Barlow *et al.*[21,22] have postulated that a highly substituted "hypervariable" loop in each SCR is responsible for specific ligand binding, although in other proteins, the heparin-binding site is spread over discontinuous regions of many amino acids dispersed in sequence space.[23] While this "hypervariable" loop contains the residues HGRK in hfH SCR 7, it alone cannot account for the entire heparin-binding site.

The fH1516 structure[21] has formed the basis for generating models of homologous complement regions in proteoglycans (C-terminal SCR domain of the G3 region),[24] C4b-binding protein (three β-chain SCRs or "CP modules"),[25] human complement protease C1s (modules IV and V of the catalytic region)[26] and the

complement receptor 2 (two SCRs of the C3 binding domain).[27] However, fH1516 has not been used to model the other fH SCR domains.

We now report specific amino acids that form the putative heparin-binding site on hfH SCR 7, based on the detailed analysis of its sequence with those of other heparin-binding homologous domains: SCR 7 from murine/bovine fH (mfH and bfH; D.A. Male *et al.*, unpublished results) and SCR 2 from human fHR-3;[16] molecular modeling of hfH SCRs 6 and 7 (fH67, using fH1516 and vcp34 as templates) and surface electrostatic potential calculations. There appear to be two primary (conserved) sites of heparin/sialic acid interaction: in the interdomain linker between SCRs 6 and 7 and at the end of the "hypervariable" loop,[21] and two secondary (sequence-specific) sites flanking the latter.

## 2. Methods

### *2.1 Sequence retrieval and alignment*

The sequences for SCR 7 of hfH, mfH and bfH, and SCR 2 of fHR-3 were retrieved from protein sequence databases (SWISS-PROT and PIR) at the Australian National Genomic Information Service (http://www.angis.org.au). Multiple sequence alignment was carried out using CLUSTALW[28] (with default BLOSUM scoring matrices for pairwise and multiple alignment); and MALIGN[29] (with the default scoring matrix derived from multiple-structure alignments). The alignments were visually edited for the alignment of conserved and chemically similar residues.

### *2.2 Model building*

The fH67 structural model is based on the alignment of its sequence with those of fH1516 and vcp34, using CLUSTALW[28] and then checked with that obtained from MALIGN.[29] A composite alignment was then constructed and revised by visual editing, to minimize gaps and correctly align the four structurally important cysteine residues as well as the glycine, proline and tyrosine residues usually conserved in each SCR. The average NMR structures for fH1516 (PDB ID: 1HFH) and vcp34 (PDB ID: 1VVC) were used as templates in model building. The program MODELLER[30] was used to generate the fH67 structural model from the structures of fH1516 and vcp34, with specific constraints to enable the formation of two disulfide bridges within each domain. The initial model was iteratively refined using in-built molecular dynamics with simulated annealing protocols, to improve the structural quality as computed by PROCHECK.[31] and the coordinates will be deposited with the Protein Databank.

*2.3 Surface electrostatic potential calculations*

Electrostatic potentials on the molecular surface of fH67 and fH1516 were computed by the finite-difference Poisson-Boltzmann method, as implemented in GRASP[32] using the simple charge model, with all histidine residues left uncharged.

## 3. Results and Discussion

*3.1 Sequence alignment*

Figure 1 shows the alignment of SCR 7 of human/murine/bovine fH and SCR 2 of fHR-3. We have included the residues from the last (fourth) conserved cysteine residue of the preceding domain (SCR 6 in the case of fH and SCR 1 in the case of fHR-3), in order to understand the role of the interdomain linker. We note that the linker in each sequence as well as fH SCR 20 (not shown), comprises exactly three residues. This is significant since those SCRs that do not bind heparin are preceded by longer linkers, with little residue conservation. The interdomain linkers in the heparin-binding SCR sequence set contain one or two conserved basic residues (Arg-387 and Lys-388 in the case of hfH) and, more importantly, no acidic ones.

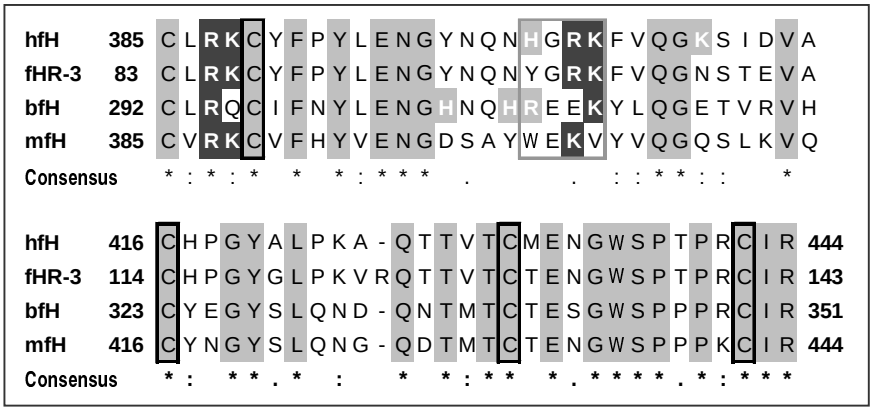


**Figure 1 Multiple sequence alignment of heparin-binding SCR domains**

CLUSTALW[28] alignment, showing conserved (*, grey backgound), conservatively substituted (:) and semi-conservatively substituted (.) positions. Conserved cysteine residues for the heparin-binding domain are in boxes. Sequence numbers correspond to the human fH (hfH: SWISSPROT CFAH_HUMAN), human fHR-3 (hfHR-3: SWISSPROT CFHD_HUMAN), bovine fH (bfH: PIR S65551) and mouse fH (mfH: SWISSPROT CFAH_MOUSE). Putative heparin-binding site residues are shown in white type with black (primary site) or grey (secondary site) background. The "hypervariable" region is enclosed by a grey box.

The end of the "hypervariable" region (Figure 1: grey box) contains conserved basic residues (Arg-404 and Lys-405 of hfH; shown in white type with black background in Figure 1), with some partly conserved basic residues at the start of this region (Figure 1: white type with grey background). The only other basic residue in the vicinity is Lys-410 of hfH, which by virtue of its long sidechain, could contribute to a heparin-binding site. However, this position is occupied by polar or acidic residues in the other sequences. The proximity of Lys-410 to the identified basic residue clusters in space needs to be evaluated. Only amino acids within seven residues of the basic cluster at the end of the "hypervariable" loop were considered, based on the experimental work of Fromm *et al.*[23] Since heparin carries several negative charges, it is possible that histidine residues at the ligand-binding surface can become positively charged under the vicinity of this nucleophilic ligand, and thus contribute favorably to heparin-binding. Histidine residues occur in heparin-binding sequences of L-type C channel, TGF β1 and apo B100[23] although to date, no functional role has been ascribed to this amino acid in heparin binding. His-402 of hfH, and His-305 and His-308 of bfH can thus be considered probable secondary sites involved in heparin binding (Figure 1: white type with grey background).

The fact that heparin-binding domains are preceded by exactly a three-residue linker with conserved basic residues and the possible role of the linker in heparin-binding have not been reported before. The extremely short linker could constrain the orientation of SCR 7 relative to its predecessor, thus making only one face of the structural domain accessible to ligands.

### *2.2 Model building*

The alignment of the fH67 sequence with those of fH1516 (PDB ID: 1HFH) and vcp34 (PDB ID: 1VVC), is shown in Figure 2. fH67 shows 25% identity to fH1516 and 28% to vcp34, from pairwise MALIGN alignments (not shown). These values are close to the 'twilight zone' (around 25% identity over an alignment length of 125) defined for homologous sequences.[33] Nevertheless, the positions of the structurally important cysteine residues as well the SCR-conserved residue set (glycine, proline and tyrosine) remain aligned, permitting the development of a three-dimensional model for fH67 based on the structures of fH1516 and vcp34. It is interesting to note that the fH67 sequence is less similar to fH1516 (derived from the same human protein) than it is to vcp34 derived from the Vaccinia virus. However, given the ancient origin of hfH,[34] it is possible that the SCR domains of hfH have diverged considerably, with only key residues remaining conserved. From Figure 2, we also note that the "hypervariable" region in SCR 7 now maps on to the residues NHGR and not HGRK as proposed by Soames *et al.*[35]

The individual SCR domains from the two available NMR structures (1HFH and 1VVC) are topologically very similar.[22] Thus, the inclusion of both structures

in model building does not provide any additional structural information[36] for the SCR domain itself. Thus, in accordance with template selection rules for comparative protein modeling,[36] either structure would suffice. However, the position of SCR 16 with respect to SCR 15 is quite variable in fH1516,[21] while both domains of vcp34 are well defined.[22] Also, the relative orientations of the two domains in fH1516 and vcp34 are very dissimiliar,[22] necessitating the inclusion of both structures as templates for model building. From Figure 2, it is worth noting that the interdomain linker in fH67 (residues 385-387 LRK) is one residue shorter than the linkers of the two template structures, so that this region is essentially reconstructed by the model building program, MODELLER.[30]

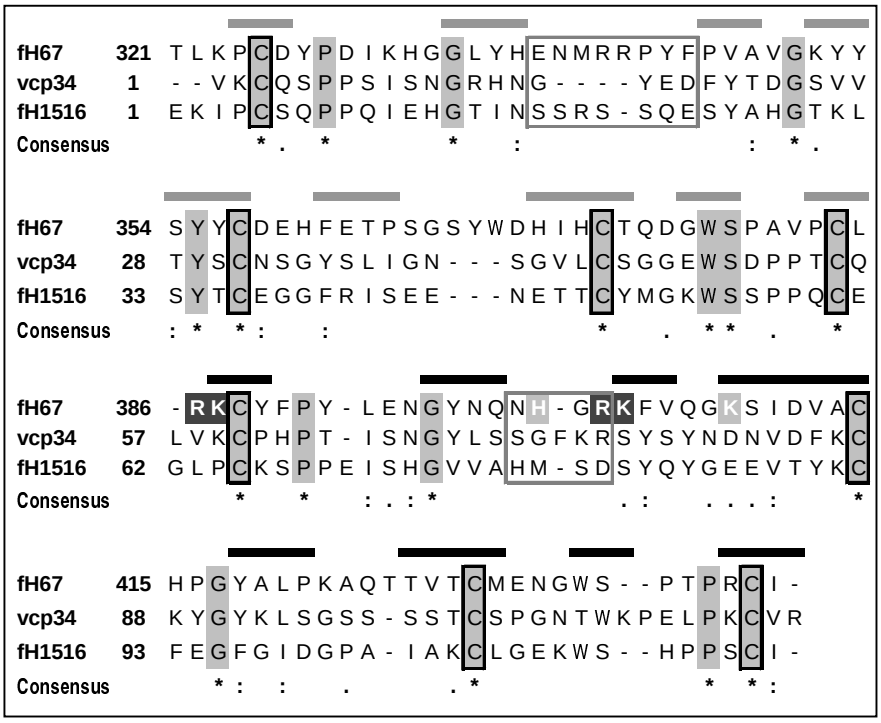


**Figure 2 Sequence alignment for model building**

Sequence numbers correspond to the human fH (fH67: SWISSPROT CFAH_HUMAN), fH1516 (PDB ID: 1HFH) and vcp34 (PDB ID: 1VVC). Horizontal bars (grey: first SCR and black: second SCR) represent β-strand regions of 1HFH and 1VVC. Sequence conservation, hypervariable loops, putative heparin-binding residues shown as in Figure 1.

The final fH67 model, following refinement as described in the **Methods** section, is shown in Figure 3(a). The overall quality of the refined fH67 model was

assessed as very satisfactory by PROCHECK,[31] with 98.0% of the residues in allowed backbone conformations. A structural overlay of the model with the template structures shows a clear resemblance to fH1516 over vcp34, particularly in the relative disposition of the two domains. The second domain of fH67 (SCR 7) is however, shifted about 2.5 Å closer to the first domain (SCR 6) and rotated 20° (away from the plane of the paper) about the linker region, compared with the two domains of fH1516. The acidic groups (carboxyl and sulfate) on the ligands that bind to SCR 7, sialic acid (PDB ID: 1NSC[38]) and the disaccharide repeat unit of the hexameric heparin (PDB ID: 1HPN[39]) are also shown in Figure 3.

Despite the fairly low sequence similarity between fH67 and fH1516, the two structures superimpose with an overall RMS deviation (RMSD) of 1.2 Å. The first domain of fH67 (SCR 6) is very similar to SCR 15 of fH1516 (RMSD of 0.91 Å over 58 Cα positions), while only 31 Cα positions of SCR 7 superimpose with SCR 16, with an RMSD of 0.85 Å.

Within each domain of fH67, the conserved disulfide bridges (SS1 and SS2 in Figure 3(a)) are placed respectively near the N- and C-terminal ends of the SCR. The disulfide bridges of fH1516 have been omitted from Figure 3(a) for reasons of clarity. The β-sheet structure of each SCR (Figure 2) is essentially retained. Since the two individual SCRs in fH67 are closer to each other than those in fH1516 or in vcp34 (with an almost extended disposition of its two SCRs), it was interesting to look for regions of interdomain interaction. There is neither any hydrogen-bonding interaction between the two domains of fH67 (results from the WHAT IF[40] server at http://www.sander.embl-heidelberg/server2/, not shown) nor hydrophobic contacts, showing their relative independence in structure and therefore, function. This lack of interaction is also noted between the two domains of fH1516, with greater interdomain separation than in the fH67 model. Since the heparin-binding SCRs form a homologous sequence set (Figure 1), the fH67 model serves as a consensus structure for SCRs 6, 7 of bfH and mfH and SCRs 1, 2 of fHR-3.

The most interesting finding of the model building exercise is the co-localization of three clusters of basic residues, on one face of the structure, accessible to solvent and ligands. The sidechains of these residues (Arg-387, Lys-388, Arg-404, Lys-405 and Lys-410) are shown in Figure 3(a). These amino acids are implicated in the heparin/sialic acid-binding site from sequence analysis, by scanning for basic clusters unique to those SCRs that bind heparin. However, their spatial disposition, could not be inferred on the basis of sequence analysis alone. The role of the interdomain linker between SCRs 6 and 7 appears to be two-fold: firstly to orient SCR 7, such that only one face of SCR 7 is accessible to ligands and secondly, to augment the density of basic residues on this face. Heparin-binding experiments to the truncated hfH containing SCRs 1-7 and with Arg-387 and Lys-388 mutated to alanine, has provided encouraging results (D.A. Male *et al.*, unpublished results) to our proposed heparin-binding site.

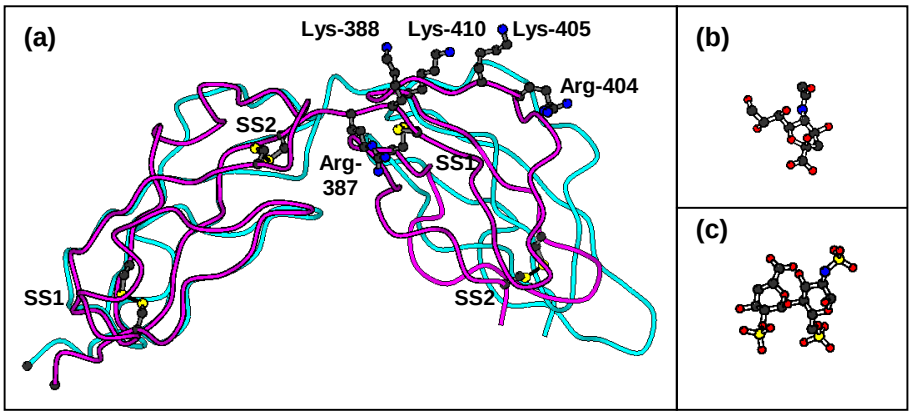


**Figure 3 Structures for the fH67 model: SCRs 6 and 7 of human fH and its ligands**

MOLSCRIPT[37] cartoon, showing (a) the Cα trace of fH67 (magenta) and fH1516 (cyan), with the respective N-termini indicated by black spheres and showing the respective first domains (SCR 6 and SCR 15; overlaid) to the left and the second domains (SCR 7 and SCR 16) to the right. Heavy atoms of the basic residues forming the putative heparin-binding site (Figures 1 and 2) and the conserved cysteine residues are shown in ball-and-stick representation (atoms colours: carbon, black; nitrogen, blue; sulfur, yellow and oxygen, red) with the two disulfide bridges (labeled SS1 and SS2) in each domain of fH67 shown as black bars. Ball-and-stick structures (colored by atom) of (b) sialic acid and (c) the disaccharide repeat unit of heparin are also shown.

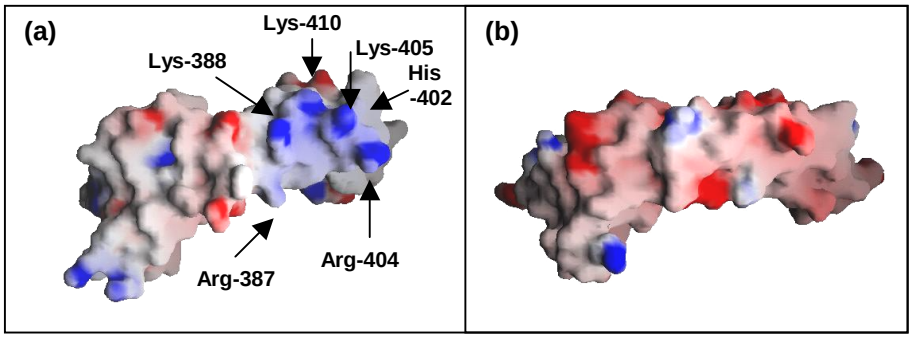


**Figure 4 Surface electrostatic potential of fH67**

Comparison of the GRASP[32] surface electrostatic potentials of (a) the fH67 model and (b) fH1516, respectively, rotated 45° about the x-axis from the view shown in Figures 3, with the interdomain linker towards the viewer; and shaded from red (< –10 kT) through white (0 kT ) to blue (> +10 kT) .

*3.3 Surface electrostatic potentials for the fH67 model*

The electrostatic potential on the molecular surface of fH67 shows both the density of charged residues on the surface (accessible *vs.* buried charged residues) as well as the character of the charged surface. It takes into account the presence or absence of oppositely charged residues next to each other and the dilution of charged surfaces by intervening uncharged residues. Figure 4 shows the surface electrostatic potential of fH67. The overall electrostatic charge on fH67 is zero, compared to $-7q$ on fH1516. While the interdomain linker in fH67 and its flanking surface from SCR 7 show positive potential (blue regions in Figure 4), the corresponding region in fH1516 shows segregated regions of negative potential (colored red in Figure 4), interspersed with zero potential. The reverse surface of fH67 (not shown) is essentially uncharged. The experimental heparin-binding by SCR 7 and not by SCR 16 support these observations.

The residues contributing significantly to the positive potential on one face of the surface of fH67 are: Lys-388, Arg-404, Lys-405 and Lys-410. While Arg-387 is at the molecular surface on this face of the model, its sidechain is almost completely extended and in this conformation (or rotamer state), its contribution to the surface potential is not appreciable. However, since Arg-387 is the only basic residue that is completely conserved in the alignment of heparin-binding SCRs (shown in Figure 1), it is possible that the rotamer observed in the fH67 model is not the functional one. Of the secondary sites suggested from sequence analysis (shown in Figure 1) for hfH, both His-402 and Lys-410 are solvent accessible and co-localized with the essential or primary 'RK' pairs. Their contribution to the heparin-binding site, particularly that of Lys-410 would be valuable.

In fHR-3, the two primary sites of hfH are conserved (Arg-85, Lys-86, Arg-102 and Lys-103) while both secondary site residues are replaced respectively by Tyr-100 and Asn-108. These residues are uncharged and will not be able to provide positive potential for heparin binding. However, polar amino acids (Gln and Asn) have been observed to stabilize the ligand-binding site, in the structures of acidic[41] and basic[42] fibroblast growth factor complexes with heparin analogs. Asn-108 is thus a possible hydrogen-bonding partner for heparin. With only the primary sites, fHR-3 provides a reduced description for the SCR 7 heparin-binding site.

Both mfH and bfH show substitutions in the primary heparin-binding sites. While murine fH has one conserved 'RK' pair (Arg-387 and Lys-388), the bovine analog has only one basic residue conserved in each cluster (Arg-294 and Lys-312). Val-405, instead of hfH Lys-405, in the mouse sequence will provide no charge or hydrogen-bonding stabilization to the ligand. However, due to its small size, the long sidechain of spatially adjacent Lys-413 (aligning with hfH Asp-413, fHR-3 Glu-111 and bfH Arg-320) could adopt a conformation favorable to heparin binding. The inclusion of mfH Lys-413 and bfH Arg-320 is supported by the

observation of partially conserved basic residues in the heparin-binding site of the Alzheimer amyloid precursor protein family.[41] The identification of mfH Lys-413 and bfH Arg-320 as possible heparin-binding residues was possible only from molecular modeling, since there is no charge conservation at this position in the alignment shown in Figure 1. The sidechain of Asp-413 is pointing away from the heparin-binding surface in the fH67 model. In mouse fH, the secondary site residues are Trp-402 and Gln-410, with the latter residue possibly forming hydrogen-binding interactions with heparin, analogous to similar interactions observed in the fibroblast growth factor structures.[42,43] Gln-295 (aligning with hfH Lys-388 in the interdomain linker) is ascribed a similar role in bovine fH, while Glu-311 (aligning with hfH Arg-404) would partially annul the positive charge of the neighboring Lys-312 in the "hypervariable" loop. However, bfH Arg-309 in one of the secondary binding sites and Arg-320 (aligning with mfH Lys-413) will provide the necessary positive potential for the experimentally observed heparin-binding. His-308 suggested as a secondary site (marked in Figure 1) for bfH from sequence alignment is facing away from the putative heparin-binding surface and is hence not part of the binding site, while His-305 (homologous to hfH Tyr-398) with a surface location would provide additional stability to heparin/sialic acid binding. The location of two histidine residues on the heparin-binding surface of the Alzheimer amyloid precursor protein crystal structure[41] further supports our hypothesis of an active heparin-binding role for this amino acid.

## Conclusions

From a combination of sequence analysis, comparative modeling and surface electrostatic potential calculations, we propose a consensus model for the heparin/ sialic acid-binding site of the human fH SCR 7 and its homologues. Conserved basic residue clusters in the interdomain linker preceding SCR 7 of hfH and at the end of the "hypervariable" loop form the primary sites of heparin interaction, with secondary sites composed of single residues (His-402 and Lys-410 of hfH) flanking the latter primary site. The possibility of locating a similar heparin-binding site in murine/bovine fH and fHR-3 has been verified. Four basic residue clusters are thus proposed as the putative hfH heparin-binding site, with uncharged mutants at these positions under experimental investigation. We also suggest that histidine residues in the vicinity of the intensely positively charged region, can become protonated, under the influence of an approaching ligand such as heparin or sialic acid.

## References

1. P.F. Zipfel, T.S. Jokiranta, J. Hellwage, V. Koistinen and S. Meri, "The factor H protein family" *Immunopharmacol.* **42**, 53 (1999)

2. K. Whaley and S. Ruddy, "Modulation of the alternative complement pathway by β1H globulin" *J. Exp. Med.* **144**, 1147 (1976)
3. J.M. Weiler, M.R. Daha, K.F. Austen and D.T. Fearon, "Control of the amplification convertase of complement by the plasma protein β1H" *Proc. Natl. Acad. Sci. USA* **73**, 3268 (1976)
4. C.J. Soames and R.B. Sim, "Interactions between human complement components factor H, factor I and C3b" *Biochem. J.* **326**, 553 (1997)
5. K. Nabil, R. Rihin, M. Jaurand, J. Vignaud, J. Ripoche, Y. Martinet and N. Martinet, "Identification of human complement factor H as a chemotactic protein for monocytes" *Biochem. J.* **326**, 377 (1997)
6. A.K. Sharma and M.K. Pangburn, "Identification of three physically and functionally distinct binding sites for C3b in human complement factor H by deletion mutagenesis" *Proc. Natl. Acad. Sci. USA* **93**, 10996 (1996)
7. B.H. Ault, B.Z. Schmidt, N.L. Fowler, C.E. Kashtan, A.E. Ahmed, B.A. Vogt and H.R. Colten, "Human factor H deficiency. Mutations in framework cysteine residues and block in H protein secretion and intracellular catabolism" *J. Biol. Chem.* **272**, 25168 (1997)
8. P.Warwicker, R.L. Donne, J.A. Goodship, T.H.J. Goodship, A.J. Howie, D.S. Kumararatne, R.A. Thompson and C.M. Taylor, "Familial relapsing haemolytic uraemic syndrome and complement factor H deficiency" *Nephrol. Dial. Transplant.* **14**, 1229 (1999)
9. M.K. Pangburn, M.A.L. Atkinson and S. Meri, "Localization of the heparin-binding site on complement factor H" *J. Biol. Chem.* **266**, 16847 (1991)
10. M.K. Pangburn, D.C. Morrison, R.D. Schreiber and H.J. Muller-Eberhard, "Activation of the alternative complement pathway: recognition of surface structures on activators by bound C3b" *J. Immunol.* **124**, 977 (1980)
11. M. Baron, D.G. Norman and I.D. Campbell, "Protein modules" *Trends Biochem. Sci.* **16**, 13 (1991)
12. K.B.M. Reid and A.J. Day, "Structure-function relationships of the complement components" *Immunol. Today* **10**, 177 (1989)
13. Y. Nakano, K. Sumida, N. Kikuta, N. Muira, T. Kobe and M. Tomita, "Complete determination of disulfide bonds localized within the short consensus repeat units of decay accelerating factor (CD55 antigen)" *Biochim. Biophys. Acta* **1116**, 235 (1992)
14. R.G. DiScipio, "Ultrastructures and interactions of complement factors H and I" *J. Immunol.* **149**, 2592 (1992)
15. T.K. Blackmore, T.A. Sadlon, H.M. Ward, D.M. Lublin and D.L. Gordon, "Identification of a heparin binding domain in the seventh short consensus repeat of complement factor H" *J. Immunol.* **157**, 5422 (1996)
16. T.K. Blackmore, J. Hellwage, T.A. Sadlon, N. Higgs, P.F. Zipfel, H.M. Ward and D.L. Gordon "Identification of the second heparin-binding domain in human complement factor H" *J. Immunol.* **160**, 3342 (1998)

17. T.K. Blackmore, V.A. Fischetti, T.A. Sadlon, H.M. Ward and D.L. Gordon, "M protein of the Group A *Streptococcus* binds to the seventh short consensus repeat of human complement factor H" *Infect. Immun.* **66**, 1427 (1998)
18. D.L. Kasper, "Bacterial capsule - old dogmas and new tricks" *J. Infect. Dis.* **153**, 5663 (1986)
19. A.K. Cheung, C.J. Parker, J. Janativa and E. Brynda, "Modulation of complement activation on hemodialysis membranes by immobilized heparin" *J. Am. Soc. Nephrol.* **2**, 1328 (1992)
20. E. Ovrum, T.E. Mollnes, E. Fosse, E.A. Holen, G. Tangen, M. Abdelnoor, M.L. Ringdal, R. Oyteseand P. Venge, "Complement and granulocyte activation in two different types of heparinized extracorporeal circuits" *J. Thoracic Cardiovasc. Surg.* **110,** 1623 (1995)
21. P.N. Barlow, A. Steinlasserer, D.G. Norman, B. Kieffer, A.P. Wiles, R.B. Sim and I.D. Campbell, "Solution structure of a pair of complement modules by nuclear magnetic resonance" *J. Mol. Biol.* 232, 268 (1993)
22. A.P. Wiles, G. Shaw, J. Bright, A. Perczel, I.D. Campbell and P.N. Barlow, "NMR studies of a viral complement protein that mimics the regulators of complement activation" *J. Mol. Biol.* **272**, 253 (1997)
23. J.R. Fromm, R.E. Hileman, E.E.O. Caldwell, J.M. Weiler and R.J. Linhardt, "Pattern and spacing of basic amino acids in heparin binding sites" *Arch. Biochem. Biophys.* **343**, 92 (1997)
24. N.C. Brissett and S.J. Perkins, "Conserved basic residues in the C-type lectin and short complement repeat domains of the G3 region of proteoglycans" *Biochemistry* **329**, 415 (1998)
25. B.O. Villoutreix, J.A. Fernandez, O. Teleman and J.H.Griffin, "Comparative modeling of the three CP modules of the β-chain of C4BP and evaluation of potential sites of interaction with protein S" *Protein Eng.* **8**, 1253 (1995)
26. V. Rossi, C.Gaboriaud, M. Lacroix, J.Ulrich, J.C.Fontecilla-Camps, J. Gagnon and G.J. Arlaud, "Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling" *Biochemistry* **34**, 7311 (1995)
27. H. Molina, S.J. Perkins, J. Guthridge, J. Gorka, T. Kinoshita and V.M. Holers, "Characterization of a complement receptor 2 (CR2, CD21) ligand binding site for C3. An initial model of ligand interaction with two linked short consensus repeat modules" *J. Immunol.* **154**, 5426 (1995)
28. J.D. Thompson, D.G. Higgins and T.J. Gibson, "Improving the sensitivity of multiple sequence alignments through sequence weighting position-specific gap penalties and weight matrix choice" *Nucleic Acids Res.* **22**, 4673 (1994)
29. M.S. Johnson, J.P. Overington and T.L. Blundell, "Alignment and searching for common protein folds using a databank of structural alignments" *J. Mol. Biol.* **231**, 735 (1993)

30. A. Sali and T.L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints" *J. Mol. Biol.* **234**, 779 (1993)
31. R.A. Laskowski, M.W. McArthur, D.S. Moss and J.M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures" *J. Appl. Crystallogr.* **26**, 283 (1993)
32. A. Nicholls, K.A. Sharp, and B. Honig, "Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons" *Proteins Struct. Funct. Genet.* **11**, 281 (1991)
33. C. Sander and R. Schneider, "Database of homology-derived protein structures and the structural meaning of sequence alignment" *Proteins Struct. Funct. Genet.* **9**, 56 (1991)
34. J. Krushkal, C. Kemper and I. Gigli, "Ancient origin of human complement factor H" *J. Mol. Evol.* **47**, 625 (1998)
35. C.J. Soames, A.J. Day and R.B. Sim, "Prediction from sequence comparisons of residues of factor H involved in the interaction with complement component C3b" *Biochem. J.* **315**, 523 (1996)
36. A. Sali, and J.P. Overington, "Derivation of rules for comparative protein modeling from a database of protein structure alignments" *Protein Sci.* **3**, 1582 (1994)
37. P.J. Kraulis, "MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures" *J. Appl. Crystallogr.* **24**, 946 (1991)
38. W.P. Burmeister, R.W. Ruigrok and S.Cusack, "The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid" *EMBO J.* **11**, 49-56 (1992)
39. B. Mulloy, M.J.Forster, C.Jones and D.B.Davies, "N.m.r. and molecular-modelling studies of the solution conformation of heparin" *Biochem. J.* **293**, 849 (1993)
40. G.Vriend, "WHAT IF: A molecular modeling and drug design program" *J. Mol. Graph.* **8**, 52 (1990)
41. J. Rossjohn, R. Cappai, S.C. Feil, A. Henry, W.J. McKinstry, D. Galatis, L. Hesse, G. Multhaup, K. Beyreuther, C.L. Masters and M.W. Parker, "Crystal structure of the N-terminal, growth factor-like domain of Alzheimer amyloid precursor protein" *Nat. Struct. Biol.* **6**, 327 (1999)
42. A. Pineda-Lucena, M.A. Jiminez, R.M. Lozano, J.L. Nieto, J. Santoro, M. Rico and G. Gimenez-Gallego, "Three-dimensional structure of acidic fibroblast growth factor in solution: effects of binding to a heparin functional analog" *J. Mol. Biol.* **264**, 162 (1996)
43. S. Faham, R.E. Hileman, J.R. Fromm, R.J. Linhardt and D.C. Rees, "Heparin structure and interactions with basic fibroblast growth factor" *Science* **271**, 1116 (1996)