

Analysis of the FGFR Signalling Network with Heparin as Co-Receptor: Evidence for the expansion of the core FGFR signaling network

Ruoyan Xu¹, Timothy R. Rudd^{1, 2}, Ashley J Hughes^{1,2}, Giuliano Siligardi^{1, 2}, David G. Fernig¹ and Edwin A. Yates¹

¹*Department of Structural and Chemical Biology, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB UK*

²*Diamond Light Source Ltd., Harwell Innovation Campus, Didcot, Oxfordshire OX11 8DE, UK*

To whom correspondence should be addressed: Edwin A Yates; Department of Structural and Chemical Biology, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB UK; eayates@liv.ac.uk; Tel: +44-(0)151-795-4429.

Running Title: *FGFR Signaling network analysis*

Keywords: *Fibroblast Growth Factor, Fibroblast Growth Factor Receptor, Heparan sulfate, glycosaminoglycans, Signaling Network*

Summary

Evolution of the fibroblast growth factor (FGF) - FGF receptor (FGFR) signaling system follows closely that of multicellular organisms. The abilities of 9 FGFs (FGFs 1-9; examples of FGF sub-families -1, -4, -7, -8 and -9) and 7 FGF receptors or isoforms (FGFR1b, 1c, 2b, 2c, 3b, 3c and 4) to support signaling in the presence of heparin, a proxy for the cellular heparan sulfate co-receptor, were assembled into a network. A connection between two FGFRs was defined as their mutual ability to signal with a particular FGF. The network contained a core of 4 receptors (FGFR1c, 2c, 3c and 4) with complete connectivity and high redundancy. Analysis of the wider network indicated that neither FGF-3 nor FGF-7 was well-connected to this core of 4 receptors and, that divergence of a precursor of FGF sub-groups 1, 4 and 9 from FGF subgroup 8 may have allowed expansion from a 3 member FGFR core signaling system to the 4 member core network. This increases 4-fold the number of possible signaling combinations. Synchrotron Radiation Circular dichroism spectra of the FGFs with heparin revealed no overall common structural change, suggesting distinct heparin binding sites throughout the FGFs. The approach provides a potential method of identifying agents capable of influencing particular FGF/FGFR combinations, or areas of the signaling network for experimental, or therapeutic purposes.

Introduction

The fibroblast growth factors (FGFs) regulate many aspects of embryonic development and adult homeostasis. The 22 human FGFs are usually divided into 7 sub-families according to their sequence similarities; subfamilies -1, -4, -7, -8, -9, -11 and -19. These correspond broadly to certain functions, although the relationships of these remains a matter of debate, though the FGF-11 subfamily is intracellular and does not interact with extracellular receptors {Itoh, 2004

{Oulion, 2012 #73}. In humans, the 22 *fgf* genes encode the ligands and 5 *fgfr* genes encode the cognate membrane receptors, FGFRs. In simpler organisms, less elaborate FGF-FGFR signaling systems are evident. For example, in *C. elegans*, only two *fgf* genes and one *fgfr* gene have been identified, indicating that the FGF ligand-receptor system expanded during the evolutionary process from primitive metazoa to vertebrates {Itoh, 2004 #72;Oulion, 2012 #73}.

In structural terms, FGFR consist of two to three extracellular immunoglobulin (Ig) domains, linked by a single transmembrane region to a cytoplasmic domain, which contains a tyrosine kinase in FGFR-1 to -4 {Itoh, 2004 #72;Trueb, 2011 #102}. The latter becomes phosphorylated following binding of a FGF to the receptor together with its obligatory co-receptor, heparan sulfate (HS), which is the carbohydrate portion of cell surface HS proteoglycans (HSPGs). Phosphorylation of the cytoplasmic domain leads to downstream signaling events. The FGFR signaling system has been implicated in both normal developmental and aberrant processes, and the possibility of controlling it remains of great interest for the study and eventual treatment of many diseases and to promote tissue repair {Beenken, 2009 #96;Grose, 2005 #86;Presta, 2005 #85}. Alternative splicing produces receptors with differences in the third Ig domain of FGFR (D3) and/or shortened versions, in which the acid box between D1 and D2 is absent. According to the crystal structures of FGFs with receptor domains D2 and D3, FGFs bind to D2 and the proximal regions of D3 and this may explain the specificity for FGFs between splice variants {Plotnikov, 2000 #77;Plotnikov, 1999 #76}. FGFRs 1 and 2 exhibit distinct binding properties to heparin {Powell, 2002 #79} and the crystal structures of complexes of FGF and FGFR in the presence of heparin-derived oligosaccharides provide evidence of direct interaction of the sugar with each protein component, albeit in two distinct arrangements {Pellegrini, 2000 #81;Schlessinger, 2000 #82}. The stimulation of cell proliferation requires the formation of a ternary complex of the FGF, FGFR and the HS co-receptor, a requirement demonstrated *in vivo* {Lin, 1999 #103}. The HS co-receptor has at least one other key function, since it controls the transport of the FGFs from source to targets {Duchesne, 2012 #39}.

HS, as well as its proxy heparin, are glycosaminoglycan (GAG) polysaccharides with a common underlying disaccharide backbone structure, comprising a uronic acid, α -L-iduronic or β -D-glucuronic acid, linked alternately through a 1-4 linkage to an α -D-glucosamine residue. Biosynthesis first generates a polysaccharide chain containing 50 to 100 repeating disaccharides of 1, 4 linked β -D-glucuronic acid and α -linked *N*-acetyl-D-glucosamine, which is then modified enzymatically in the Golgi, first through *N*-deacetylation and *N*-sulfation by *N*-sulfotransferases acting concomitantly to produce clusters of *N*-sulfated glucosamines in the chain. These provide sites for the action of subsequent enzymes, including the epimerization of D-glucuronate acid to L-iduronate acid, *O*-sulfation of C-2 on iduronate residues and of C-6 and C-3 on glucosamine residues {Ori, 2008 #28}. The result is that modifications tend to be clustered and resemble distinct domains of varying dimensions known respectively as; non-sulfated (*N*-acetylated) domains, intermediate domains (NAS) domains, that have approximately one in two D-glucosamine units modified by *N*-sulfation, and sulfated domains, in which glucosamine is *N*-sulfated (termed S-domains). A branched arrangement of this biosynthetic pathway, based on a disaccharide unit and commensurate with the abundances of experimentally observed structures, has been suggested recently {Rudd, 2012 #95}. The more highly sulfated regions of HS resemble heparin, but contain fewer *O*-sulfate groups, and are generally considered to provide protein binding sites {Ori, 2008 #28}.

The interactions of some FGFs with HS and its derivatives have been subjected to scrutiny *in vitro*, but their levels of specificity and selectivity are actively debated, e.g., {Kreuger, 2006 #104}. A recent biophysical analysis suggested that different FGFs bind to heparin in distinct ways to different, yet overlapping motifs in the polysaccharide {Xu, 2012 #47}. It has been proposed that a higher degree of specificity may exist at the level of the full ligand-receptor system {Guimond, 2006 #71; Kreuger, 2006 #104}. *In vitro*, some specificity is evident, in particular regarding the preferences of FGFR for FGFs. However, the majority of FGF sub-families can support signaling through several receptors in the presence of heparin, as a proxy for HS {Guimond, 2006 #71; Ornitz, 1996 #64}. This is likely to identify all the possible signaling combinations because, generally, the extensively sulfated heparin activates all pathways in which only a restricted set of forms of HS are active. *In vivo* ternary complexes containing the FGF, HS and FGFR can be formed *in situ* only on HS polysaccharide co-receptors from particular tissue compartments {Allen, 2003 #105}. Thus, a higher level of selectivity is likely *in vivo*, owing to the lower sulfation and restricted sets of structures in HS.

The expansion of FGFs and FGFRs during evolution and the pivotal regulatory role played by the HS co-receptor suggests that specificity may be discernible from a consideration of the entire system rather than individual interactions. Thus, we have analysed the properties of the FGF-FGFR-HS signaling system from the point of view of a simple network, in which the ability of a particular FGF to support signaling is represented as connecting any two or more FGFRs through which it can signal. Alongside, we have measured the changes in the secondary structure of selected FGFs that accompany their interaction with heparin to determine if these follow a simple pattern.

Results

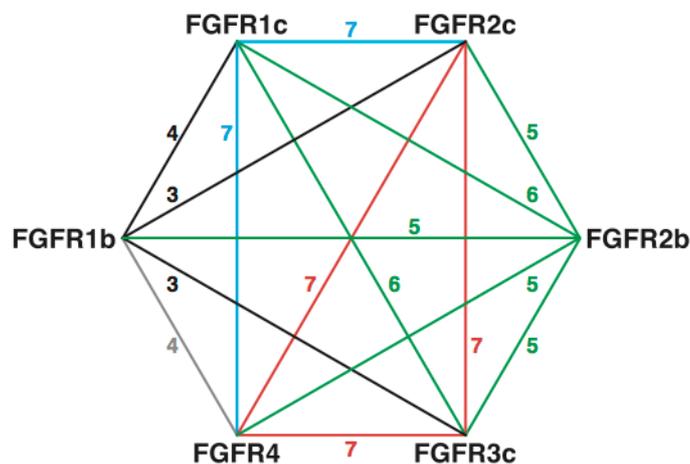


Figure 1A. A representation of the signaling network between FGF receptors. The numbers on the edges represent the number of FGFs that are able to support signaling between the respective FGFRs in the presence of heparin. The products of the published mitogenic activities are also higher in the core network formed by FGFR1c, 2c, 3c and 4, than elsewhere.

(i) **Analysis of the FGFR signaling network.** One important characteristic of the FGF-FGFR signaling system is its high degree of connectivity. Thus, in many cases, different FGFs are known to have the ability to interact with, and signal through, several FGFRs. The interactions of FGFs

with FGFRs have been surveyed previously in the presence of heparin in terms of both binding and the ability to support signaling in a cell-based assay {Ornitz, 1996 #64}. Binding events *per se* do not necessarily equate to the formation of a signaling complex, and so the data reported for signaling in an experimental cell-based assay of mitogenesis for FGF-1 to -9 and with receptors FGFR1b, 1c, 2b, 2c, 3c and 4, rather than binding were analysed. This allows a preliminary network to be assembled, employing approaches from discrete mathematics (graph theory), in which two FGFRs (represented by *vertices*) that both signal with a particular FGF are shown joined (by *edges*) (represented by **Table 1**). The ensemble of these “connections”, records the ability of the FGF to affect multiple points in a network and can be analysed in terms of both the individual FGFs and the receptors through which they signal in the presence of heparin but, more interestingly, also form a graph depicting connectivity between receptors (**Figure 1A**). This could represent, for example, a signaling situation that may exist between two cells supported by the same FGF (and HS structures), which are expressing distinct receptors, or combinations of receptors. It is evident that some receptors, for example, FGFR2c are more heavily connected than others, such as FGFR2b and the level of redundancy is noteworthy. It can also be seen that there is a highly connected core network, involving FGFR1c, 2c, 3c and 4, which are all interconnected and the connections are heavily redundant (**Figure 1A**). The network generated using this approach lends itself to analysis with tools from Graph Theory {Van Mieghem, 2010 #99}, which demonstrates that this is a substantially connected network (**Table 1**). It is known, however, that beyond this immediate FGFR network, HS interacts with over 400 proteins {Ori, 2011 #23} but, there is little experimental data concerning those interactions and any relation they may have with FGF signaling.

Table 1. Relative mitogenic activity of individual exogenous FGFs on BaF3 cells lacking HS and in the presence of added heparin (2 µg/ml). Data are from {Ornitz, 1996 #64}, with the exception of the value for FGFR1c and FGF-8, which is from {Zhang, 2006 #183}, and are normalised relative to FGF-1 and were rounded to the nearest tenth. The FGF sub-family to which each FGF has been attributed, is also shown. An expander is a graph in which each vertex (or node) is connected to each other vertex to produce a network with the same number of connections (edges) emanating from each. Such graphs are extremely robust since breaking individual connections has little effect on their overall characteristics {Van Mieghem, 2010 #99}. Eigenvalue analysis of the matrices representing the full and core networks can provide a measure of their qualities, in particular, the difference between the first and second eigenvalues, the eigenvalue gap. Analysis of the product of the matrix shown in the Table with its transpose reveals a difference in the moduli of the first and second eigen values of (17.6-2.1=15.5), which indicates a substantially connected graph {Hoory, 2006 #1}.

	FGF-n								
	1	2	3	4	5	6	7	8	9
Sub-family	1	1	3	4	4	4	7	8	9
FGFR									
1b	1.0	0.6	0.3	0.2	0.0	0.0	0.1	0.0	0.0
1c	1.0	1.0	0.0	1.0	0.6	0.6	0.0	0.6	0.2
2b	1.0	0.1	0.5	0.1	0.0	0.1	0.8	0.0	0.1
2c	1.0	0.6	0.0	0.9	0.3	0.6	0.0	0.2	0.9
3b	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4
3c	1.0	1.1	0.0	0.7	0.1	0.1	0.0	0.4	1.0
4	1.0	1.1	0.1	1.1	0.1	0.8	0.0	0.8	0.8

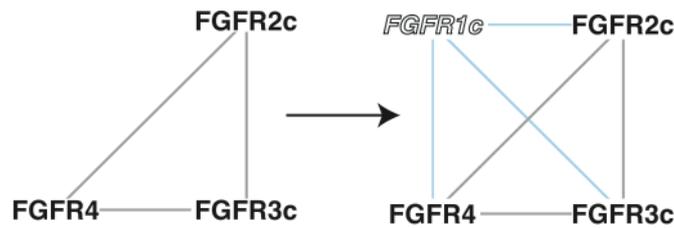


Figure 1B. The highly connected core network involving FGFR1c, 2c, 3c and 4. The core network can be differentiated into an earlier triangular sub-network (left) comprising FGFR2c, 3c and 4 which, upon divergence of FGF subfamilies 1, 4 and 9 from the FGF-8 sub-family, and evolution of FGR1c, formed a 4-member core network (right).

(ii) **Analysis of mitogenic signaling activity of a network comprising fgfs-1 to 9 with FGF receptors of the core network.** The data reporting the relative mitogenic activity of individual FGFs with FGFRs {Ornitz, 1996 #64} (**Table 1**) can be analysed by asking which FGFs can individually support signalling through pairs of FGFRs, and so form a reciprocating signaling network. This was done by taking the products of the two mitogenic activities between the FGFRs and individual FGFs (**Table 2**). In this sense, it is clear that FGFR1b, 2b and 3b are all rather poorly connected to other FGFR members, and even when they do support signaling, it is often to a relatively low extent. Notably, FGF-3 and -7 signal almost exclusively through these FGFRs, forming a sub-network, that is able to utilise FGF-3 and FGF-7, which most of the other FGFRs are unable to do (with the exception of FGFR4 with FGF-3, but then only weakly). Although FGF-3 has been reported to bind FGFR4, no mitogenic activity was reported {Ornitz, 2001 #62}.

Table 2. A measure of signaling between FGFRs in the core network can be found from the product of the mitogenic activity of the two FGFRs (from **Table 1**) and the particular FGF. This analysis assumes that there is no direct influence exerted on signaling through a FGFR by any other FGFR.

FGFR	FGF-n								
	1	2	3	4	5	6	7	8	9
1c and 2c	1.00	0.60	0.00	0.90	0.18	0.36	0.00	0.12	0.18
2c and 3c	1.00	0.60	0.00	0.81	0.03	0.06	0.00	0.08	0.90
3c and 4	1.00	1.21	0.00	0.77	0.01	0.08	0.00	0.32	0.80
1c and 4	1.00	1.10	0.00	1.10	0.06	0.48	0.00	0.48	0.16
2c and 4	1.00	0.66	0.00	0.99	0.03	0.48	0.00	0.16	0.72
1c and 3c	1.00	1.10	0.00	1.10	0.00	0.48	0.00	0.24	0.16

Exclusion of poorly connected FGFR1b, 2b and 3b, from **Table 2**, reveals a core network (**Figure 1A**) with complete mutual interconnection (*a complete graph*) and redundancy. The significance of a complete graph in a signaling network is that it is extremely robust, and high redundancy further ensures the fidelity of the signals in this section of the network. These four FGFRs are generally more highly connected in terms of the number of FGFs which support signaling between FGFR pairs, averaging 6.8 inside the core network and 4.2 outside. Furthermore, the sum of their mitogenic products (taken from **Table 2**) inside the core network is ~3.8 and outside is ~1.6 (ignoring crossed-terms).

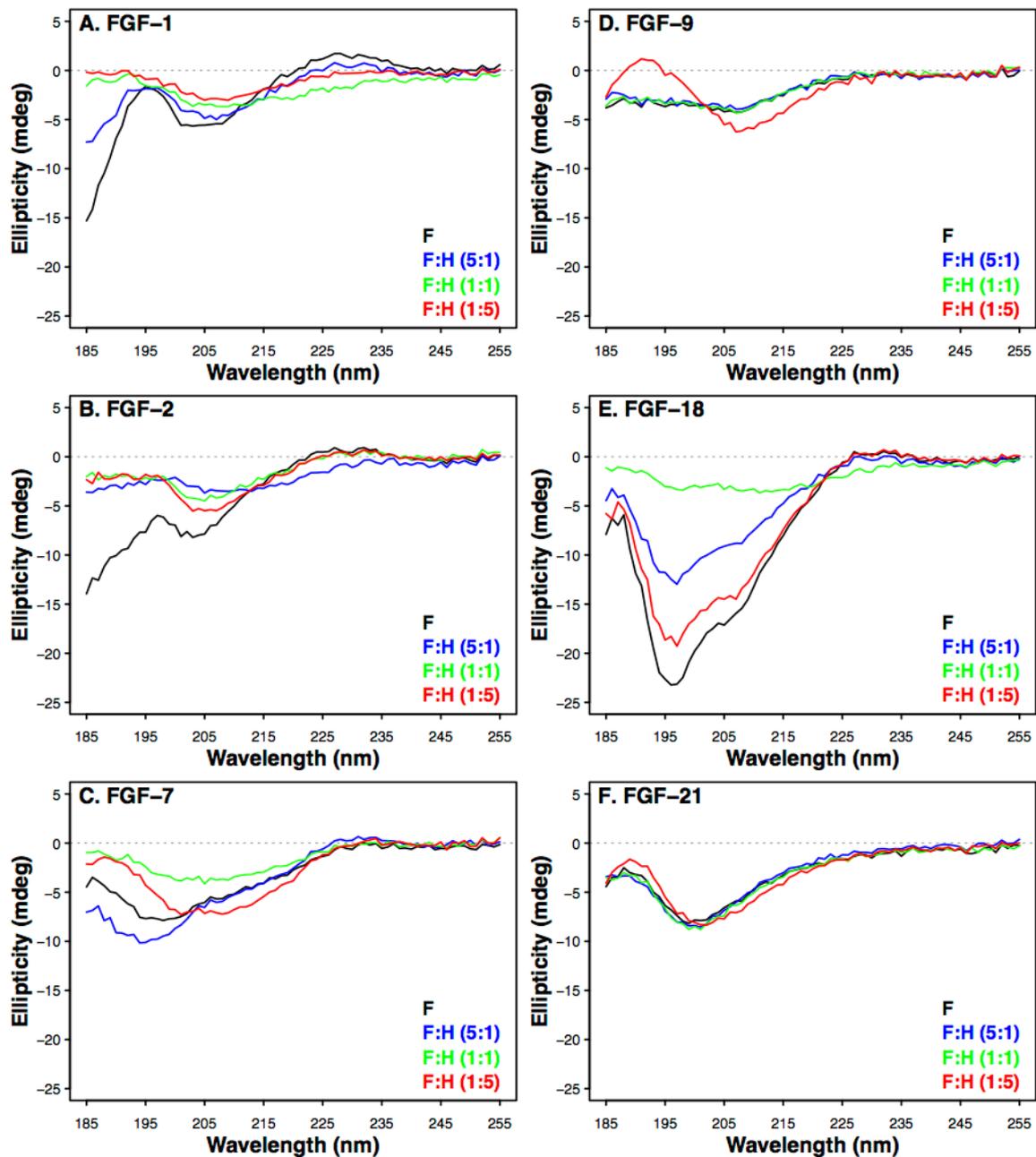


Figure 2 SRCD spectra of FGFs with different concentrations of heparin. Six FGFs and their heparin complexes spectra at molar ratios (FGF: heparin), 5:1, 1:1 and 1:5. **A**, FGF-1. **B**, FGF-2. **C**, FGF-7. **D**, FGF-9. **E**, FGF-18 (FGF-8 sub-family). **F**, FGF-21 (FGF-19 sub-family). Spectra for the FGF alone and for the 1:5 molar ratio of FGF:heparin were originally published in {Xu, 2012 #47} © the American Society for Biochemistry and Molecular Biology.

Subsequent to the work of Ornitz *et al.*, {Ornitz, 1996 #64}, Zhang *et al.*, {Zhang, 2006 #183} found that FGF-8 had mitogenic activity through FGFR1c [Table 1A] and this work also included additional members of the FGF-8 sub-family. While there is inevitably some variation in cell proliferation assay data between these two reports, it is interesting that the data reported in {Zhang, 2006 #183} for the three members of the FGF-8 subfamily (FGFs -8, -17 and -18) reveal very weak mitogenic activity for FGF-18 (<5% that of FGF-1) with FGFR1c, or weak to moderate activities for FGF-17 (23% of FGF-1) and FGF-8 (58% of FGF-1) with FGFR1c. Consequently, the products of mitogenic activities for combinations of FGFRs involving FGFR1c are the weakest 3

(with FGF-18), and 3 of the 4 weakest for both FGF-8 and FGF-17. This indicates relatively poor connectivity of FGFR1c *via* the other members of the FGF-8 subfamily, but particularly FGF-18, to FGFRs 2c, 3c and 4.

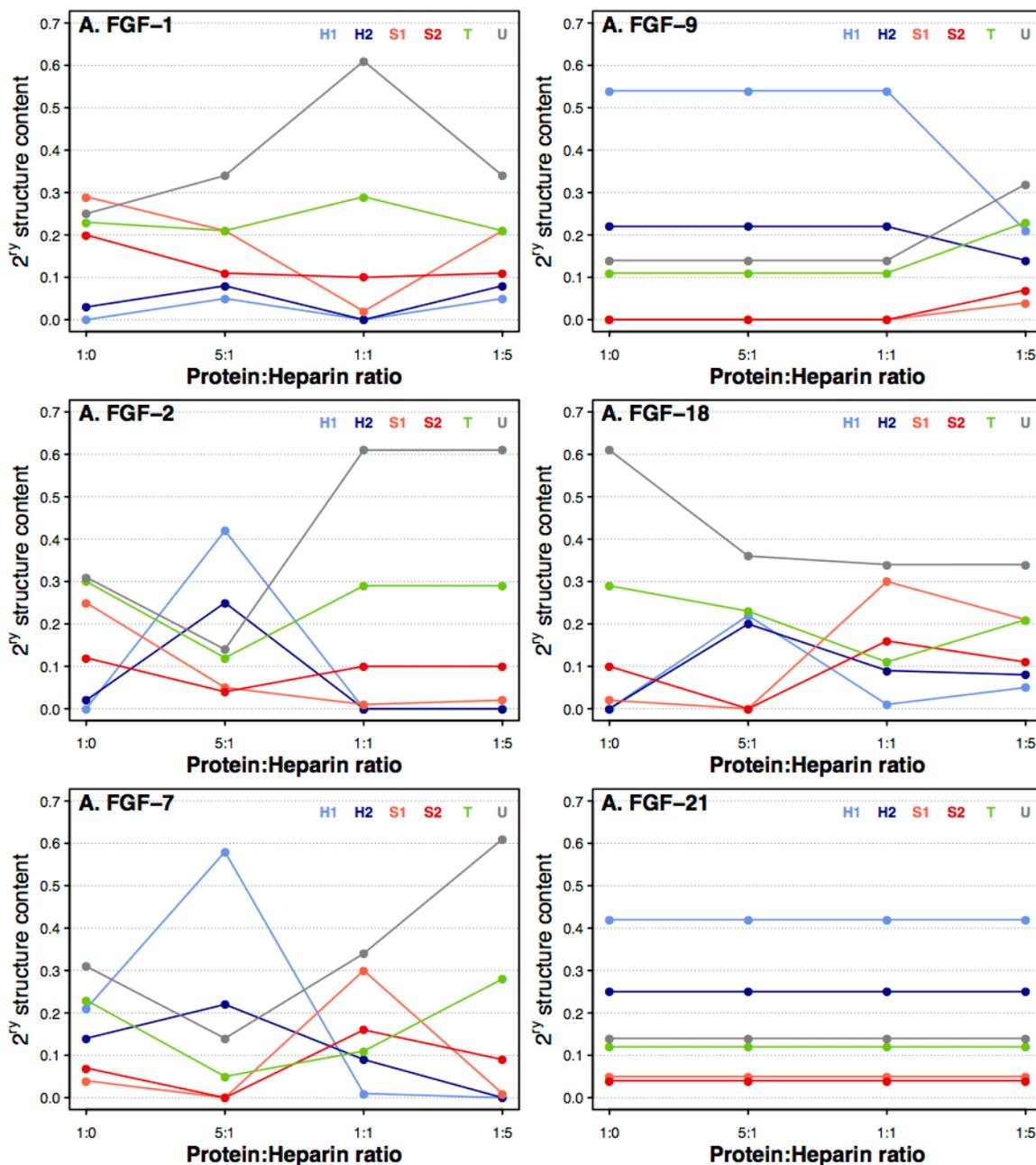


Figure 3 Secondary structure changes of FGFs with different concentrations of heparin. The secondary structures were determined for six FGFs and their heparin complexes at molar ratios (FGF: heparin), 5:1, 1:1 and 1:5. **A**, FGF-1. **B**, FGF-2. **C**, FGF-7. **D**, FGF-9. **E**, FGF-18 (FGF-8 sub-family). **F**, FGF-21 (FGF-19 sub-family). Protein secondary structures were determined using the Dichroweb website, employing the SELCON3 algorithm with reference data set 3 [45, 46]. The FGF and FGF:heparin (1:5) complex secondary structures and SRCD spectra were originally published in {Xu, 2012 #47} © the American Society for Biochemistry and Molecular Biology.

Interestingly, this implies two distinct regions within the core network. The connections between FGFR2c, 3c and 4 all involve FGF-1, -2 (subfamily 1), -4, -6 (subfamily 4), -8 (subfamily

8) and -9 (subfamily 9), while those to FGFR1c from all other FGFRs involve FGF-8 subfamily members much more weakly (**Figure 1a and 1b**). This suggests that divergence from this FGF sub-family by a common ancestor of FGFs -1, -4, -6 and -9 coupled to, or followed by co-evolution of FGFR1c (*followed by*, because FGF evolution is thought to have preceded FGFR evolution {Itoh, 2004 #72}) was a key event in allowing expansion of an original three member, triangular core network with three interconnections (**Figure 1B**) to a four member form, with six interconnections between FGFRs (**Figure 1A**). This provides a 4-fold increase in the number of possible signaling combinations (single, double, triple and quadruple). The robust four member form of the core network has a level of redundancy in each connection of 6 or 7 (from **Table 1**).

(iii) SRCD spectral changes when adding different concentrations of heparin may relate to the number of heparin binding sites of the FGFs. In light of the above, it was of interest to examine the binary complexes of FGFs and heparin. The SRCD spectra of six FGFs and their heparin complexes were collected under identical conditions (**Figure 2 A-F**). The spectrum of FGF-18 (belonging to FGF sub-family 8) (**Figure 2 E**) exhibited significant changes when heparin was introduced, while FGF-1, -2, -7 and -9 also showed some differences (**Figure 2 A-D**). On the other hand, FGF-21 (sub-family 19), which binds heparin only very weakly {Xu, 2012 #47}, showed very little spectral change (**Figure 2 F**), confirming the lack of binding and showing that higher levels of heparin do not unduly alter the CD spectrum.

Importantly, the SRCD spectra demonstrated that there was no simple trend in the structural changes observed in the FGFs as the concentration of heparin was increased (i.e. there was not a single direction of change). This is illustrated in **Figure S1**, where the normalised ellipticity at 203 nm is plotted for FGF-1, -2 and -18. Thus, although structural changes depend on the presence of heparin, higher concentrations do not necessarily produce larger changes in secondary structure in a linear direction (**Figure 4 and Table S1.**). One trend that is apparent in the secondary structure data is evident for FGF-1, -2, -7 and -9. As the heparin concentration increases, so does the content of unordered secondary structure (**Figure 4**), the opposite occurs for FGF-18; the protein becoming markedly more structured (both strands and helices), while for FGF-21 there is no change presumably because it does not interact with heparin. This may be related to the number of binding sites for the sugar in the proteins {Ori, 2009 #26; Xu, 2012 #47}. When a small amount of heparin is present, most of it will bind to the canonical heparin binding site (HBS-1), since this has the highest affinity, initiating a particular set of conformational changes in the FGF ligands. As more heparin is added, the canonical heparin binding sites become saturated and the heparin will start binding the lower affinity HBS-2, HBS-3, and HBS-4, producing additional conformational changes. Whereas the canonical HBS1 is the one engaged with the sugar in ternary receptor complexes {Pellegrini, 2000 #81; Schlessinger, 2000 #82}, the secondary binding sites may have other functions, including regulation of the transport of FGFs through extracellular matrices {Duchesne, 2012 #39}. Therefore, the specificity and effects on FGF structure of heparin binding may be more convoluted than that simply required for forming a signaling complex, since it may reflect different functions, e.g., signaling and transport.

Discussion

The present work indicates that the co-evolution of FGFR1c and divergence of FGF sub-families -1, -4 and -9 from FGF subfamily 8 may be linked to expansion of a triangular core network

involving FGFR2c, 3c and 4 into a four member network. This core network is highly redundant and this presents cells with the possibility of reliably stimulating the highly connected core network with particular FGFs in distinct tissues, while maintaining the fidelity of core signaling events. In graph terminology, this core network is a *complete* graph. Furthermore, it has good expander properties (**Table 1**); that is to say, each vertex is connected to each other vertex, and several times over, to produce a very robust system.

Members of FGF subfamily 8 and the other FGFs diverged early in the evolution of multicellular life. The analysis suggests that expansion of the three to a four member core network occurred relatively early in evolution, at the stage of *fgf* expansion. Two-stages of *fgf* gene evolution have been postulated. The first involves *fgf* expansion in the early metazoan, from 3 to 6 genes by duplication. The second, in early vertebrates, involved large-scale genome duplications. This is in contrast to FGFR family expansion, which occurred during the second phase only, but then achieved additional diversity through splice variants {Itoh, 2004 #72}. The analysis by Oulion *et al.*, {Oulion, 2012 #73}, based on a systematic analysis of recently discovered gene sequences, proposes a re-classification of the *fgf* gene family into 8, rather than 7 sub-families, the notable difference being the allocation of FGF-3 to its own sub-family, in contrast to its usual allocation to sub-family-7. It is also interesting that the same analysis indicates that genes encoding the FGF-1 sub-family (FGF-1 and-2) and FGF-8 sub-family members (FGF-8, -17, -18 and -24) were already present in the eumetazoan ancestor, indicating divergence of the FGF-8 subfamily from other sub-groups at an early stage in evolution. Unlike GAG biosynthetic enzymes, *fgf*-like genes have not been identified in the unicellular choanoflagellates, postulated to represent the descendants of the last common unicellular ancestor of metazoa {King, 2008 #106;Ori, 2011 #23}. However, two *fgf*-like genes are present in *C. elegans* (*egl-17* and *let-756*) {Burdine, 1997 #45;Roubin, 1999 #43} and three in *Drosophila*; *branchless*, *pyramus* and *thisbe* {Gryzik, 2004 #40;Stathopoulos, 2004 #41}, supporting their primordial origin.

The early divergence of FGF subfamily-8 from the other FGF subfamilies has been revealed by analysis of amino acid sequences {Ornitz, 2001 #62} and, in light of the indications here of the importance of the divergence of FGF sub-family-8 for expansion of the core FGFR signalling network from three to four members, it is interesting that the FGFs identified in simple multicellular organisms resemble particular FGFs {Birnbaum, 2005 #42}. In *C.elegans*, the 2 FGFs, *egl-17* and *let-756* resemble FGF-8 and -9 respectively, while in *Drosophila*, both *pyramus* and *thisbe* resemble FGF-8.

Heparin binding sites on FGFs overlap with FGFR binding sites. FGFs mediate their bioactivities by binding to their cell surface receptors, FGFR and HS, to form a signalling complex. It seems likely that, *in vivo*, FGF binds first to HS and then to the FGFR, since HS binding sites are several orders of magnitude more abundant in the pericellular matrix and on the cell surface {Duchesne, 2012 #39}. There are different numbers of HBSs evident in the FGF sub-families {Xu, 2012 #47}. A question of interest is whether there is any overlap between these and the FGFR binding sites. FGF-2 has three HBSs {Baird, 1988 #38;Faham, 1996 #37;Kinsella, 1998 #35;Ori, 2009 #26;Thompson, 1994 #36} and the residues interacting with the FGFR have been identified by site-directed mutagenesis {Springer, 1994 #34} and X-ray crystallography (FGF-2-FGFR1, PDB: 1CVS) {Plotnikov, 1999 #76}. A comparison of these binding sites indicates that HBS-3, towards the N-terminus, overlaps with the FGFR binding site at position K30. In addition to FGF-2,

other FGFs also have residues involved in both interactions. According to a recent analysis using a protect and label approach (Figure 2 in {Xu, 2012 #47}) and the crystal structures of heparin-FGF-1-FGFR2 (PDB: 1E00) {Pellegrini, 2000 #81}, K24 in HBS-3 overlaps with its FGFR. K-24 in FGF-1 is at the equivalent position to K30 in FGF-2 when the secondary structures of these proteins are aligned {Xu, 2012 #47}. By sequence alignment there is a possible HBS-3 at the N-terminus of FGF-7, R65, but, since this only contains arginines, it was not identified by protect and label {Xu, 2012 #47}. The sequence alignment of the FGF-7 subfamily and the FGF-10-FGFR2b crystal structure (PDB: 1NUN) {Yeh, 2003 #33}, shows that R78 of FGF-10 (allocated to FGF sub-family 7) is at the equivalent position to R65 in FGF-7 and so might be part of a HBS-3 {Xu, 2012 #47}, which is involved in binding to FGFR {Yeh, 2003 #33}. These secondary HBS are likely to impact on the transport of FGFs in extracellular matrix {Duchesne, 2012 #39}. The overlap of one of these, HBS3 with part of the FGFR binding site may imply that a re-arrangement of FGF-sugar interactions is required in some instances to allow the ternary signaling complex to form.

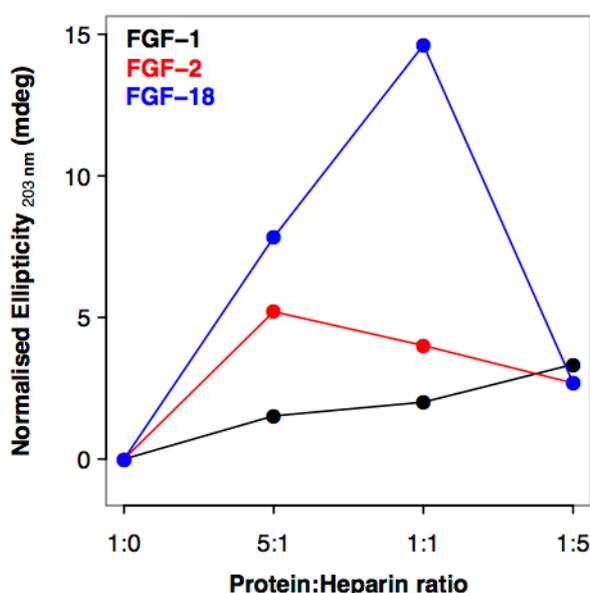


Figure 4. Normalised change in ellipticity at 203 nm for FGF bound to different ratios of heparin. This figure illustrates that the interaction with heparin and the protein does not follow a simple linear relationship. The FGF and FGF:heparin (1:5) complex secondary structures and SRCD spectra were originally published in {Xu, 2012 #47} © the American Society for Biochemistry and Molecular Biology.

The interactions between FGFs and heparin/HS derivatives and analogues depend on both their dimensions and the nature of their charge distribution, and evidence from a variety of experimental methods demonstrates that charge, size and conformation are important in these interactions {Afratis, 2012 #130; Militopoulou, 2003 #131}. Importantly, different sequences can, if the dimensions, shape and charge distribution are correct, act in the same way as heparin to provide a suitable co-receptor for signaling {Rudd, 2010 #189}. A consequence may be the observed modulatory effect of chondroitin sulfate on FGF signaling Nikitovic et al Int J Biochem Cell Biol. 2008. Interestingly, the mechanism for FGF-1 and FGF-2 have been found to be different; FGF-1 requiring only a saccharide capable of providing stabilisation while, in contrast, FGF-2 required induction of the appropriate secondary structural changes before being able to support signaling {Rudd, 2010 #189; Xu, 2012 #47}. The SRCD spectra [Figure 2 and Figure 3] undergo

considerable conformational change on addition of heparin, confirming that interactions occur and reveal differences between FGFs. Importantly, non-linear changes are evident as the ratio of heparin is increased [Figure 4]. This has been studied in some detail elsewhere and indicates the presence of several binding sites of different affinity, which have also been mapped using a 'Protect and Label' strategy {Xu, 2012 #47}.

Appreciation of the networked signaling system will influence the choice of pharmaceutical targets and the nature of attempts to target this system. Assembling the signaling network in the present way suggests that it can be approached as an entire experimental system. There are increasing indications that the combinations of activities supported by particular sugar structures may be one way of influencing this and similar networks in a more effective way than targeting single interactions e.g., {Groah, 2011 #108;Lallam-Laroye, 2011 #107;Rai, 2011 #109}. This might be particularly relevant, because it is apparent that there is a good deal of redundancy in terms of FGF signaling and sugar co-receptor specificity and selectivity. Employing the network to search for agents capable of particular combinations of selective intervention is an attractive experimental and practical possibility, as is the prospect of supporting signaling patterns distinct from those present in the extant system.

The FGF/FGFR/HS signaling network presented here is only a small proportion of a much richer, more complex network, with which HS interacts. There will also be comparable (but probably subtly different) networks of FGF/FGFR connections with other HS structures and the variability of HS structures offers an additional dimension through which the network can be viewed or influenced. Presumably, HS structures with lower sulfation levels will tend to induce restricted networks, although this remains largely untested. It will also be interesting to compare heparin binding sites and their affinities to those of different HS structures.

Highly complex signaling networks exist at the heart of cell signalling and both the FGF/FGFR system and HS are key players, interacting with many hundreds of other components. Such systems can be prone to noise, as well as being sensitive to small fluctuations causing wider disruption. It is important, therefore, for the network to have built-in stability against such sensitivity. If the core, highly connected network is viewed as serving to transmit a particular range of signals, rather than any individual ones, it is closely analogous to a communication network and can be termed an *expander graph* {Hoory, 2006 #1}, a term coined in the field of telecommunications, which describes the degree to which a network is inter-connected and reflects its resulting degree of stability. Such expanders conserve the fidelity of patterns of information transfer against noise, in this context, ensuring that the correct set of *coordinated* signals is transmitted. Consequently, the redundancy noted above is likely to reflect the need for such robustness.

Clearly, more refined data, in which the strength of signalling with other interacting proteins {Ori, 2008 #28} is included, will be required for a fuller treatment. A more complete description of this signaling system will need to incorporate the affinity of each component and take into account multiple, and potentially overlapping, binding sites. Future refinements will entail extension of the network to incorporate other proteins that interact with FGFs and FGFRs either directly, e.g., {Murakami, 2008 #20}, or indirectly through HS. Importantly, more extensive networks would provide data concerning which other receptors may be influenced and the likely capacity of the system to provide compensatory signaling pathways. A probable outcome of this approach will be the realisation that it will be necessary to inhibit (or stimulate) several events simultaneously and

the network does provide an experimental framework for screening compounds to this end. This would increase the discovery rate of such compounds, which have intriguing activities in animal models and patents, but for which there are just a few examples {Groah, 2011 #108;Lallam-Laroye, 2011 #107;Rai, 2011 #109}.

Experimental procedures

FGF expression and purification:- FGF-1 (UniProt Accession: P05230; residues: 16-155) and FGF-2 (UniProt Accession: P09038-2; residues: 1-155) were cloned into a vector pET-14b (Novagen, Merck Chemical Ltd, Nottingham, UK) and purified as described {Ke, 1992 #15;Uniewicz, 2010 #12}. cDNAs encoding FGF-7 (Uniprot Accession: P21781; residues: 32-194), FGF-9 (Uniprot Accession: P31371; residues: 1-208), FGF-18 (zFGF5) (UniProt Accession: O76093; residues: 28-207) and FGF-21 (Uniprot Accession: Q9NSA1; residues: 29-209) were inserted separately into a modified pET-24b vector (pETM-11, kind gift of Dr. Paul Elliott, University of Liverpool), and they were purified as described {Ke, 1992 #15;Uniewicz, 2010 #12;Xu, 2012 #47}.

Synchrotron radiation circular dichroism (SRCD) spectroscopy:- FGFs were buffer exchanged into CD buffer (15.3 mM Na₂HPO₄, 2.2 mM NaH₂PO₄, pH 7.5). Six different FGFs (FGF-1, -2, -7, -9, -18 and -21) at a concentration of 0.5 mg/ml (FGF-1, -2, -7, -18 and -21) or 1 mg/mL (FGF-9) and heparin (17 kDa average molecular weight, Celsus Lab, Cincinnati, OH, USA: disaccharide composition (%): ΔUA-GlcNAc; 5.8, ΔUA-GlcNAc; 8.4, ΔUA-GlcNS; 2.1, ΔUA-GlcNS(6S); 22.9, ΔUA(2S)-GlcNS; 9.0, ΔUA(2S)-GlcNS(6S); 49.5, ΔUA(2S)-GlcNAc; 0.6, ΔUA(2S)-GlcNAc(6S),1.7 (data from {Skidmore, 2006 #153}) were mixed in different ratios (1:5-5:1) and then loaded into a quartz cuvette (Hellma UK Ltd, Southend on Sea, Essex, UK) path length 0.2 mm, and the SRCD spectra were acquired from 178 nm to 260 nm on beam line B23, Diamond Light Source Ltd. {Hussain, 2012 #100}, as described {Xu, 2012 #47}. Heparin binds FGF-21 (sub-family 19) only very weakly and showed very little spectral change (**Figure 2 F**) on addition of heparin at any of the concentrations tested. This confirms both the lack of binding and that higher levels of heparin do not provide significant CD spectral features in this experimental set-up. Protein secondary structures were determined using the Dichroweb website, employing the SELCON3 algorithm with reference data set 3 {Whitmore, 2004 #118;Whitmore, 2008 #113}.

Abbreviations

ΔUA(2S); 4,5 unsaturated, 2-O-sulfated uronate residue. ΔUA; 4,5 unsaturated uronate residue. FGF; fibroblast growth factor. *fgf*; the gene encoding a fibroblast growth factor. FGFR; fibroblast growth factor receptor. GAG; glycosaminoglycan. HBS; heparin binding site. HS; heparan sulfate. HSPG; heparan sulfate proteoglycan. GlcNAc; N-acetyl glucosamine. GlcNAc(6S); N-acetyl glucosamine, 6-O-sulfate. GlcNS; N-sulfamino glucosamine. GlcNS(6S); N-sulfamino, 6-O-sulfated glucosamine. NAS-domain; intermediate domain of heparan sulfate. S-domain; sulfated domain of heparan sulfate. SRCD; synchrotron radiation circular dichroism.

Acknowledgements

The authors acknowledge the support of the Cancer and Polio Research Fund and the North West Cancer Research Fund.