UNIVERSITY OF LIVERPOOL

Institute of Integrative Biology

Specificities of the interaction of fibroblast growth factor and heparan sulfate

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of doctor in Philosophy

Yong Li

September 2015

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the regulations of the University of Liverpool. The work described is original and has not been submitted for any other degree. All aspects of the experimental design and planning for the study were conducted by me in conjunction with my supervisors, Dr. M Wilkinson and Professor D Fernig. The experimental work in this dissertation has been undertaken by me, with specific contributions that I have indicated below and in the text.

Dr. Ed Yates provided a library of chemical modified heparin. Mr. Changye Sun assisted me in the production of fibroblast growth factors and cell culture. Dr. Igor Barsukov and Dr. Barbara Franke aided with experiments performed on SEC-MALLS.

Any views in this thesis are those of the author and in no way represent those of the University of Liverpool. This thesis has not been presented to any other university for examination in the United Kingdom or overseas.

Acknowledgements

I would like to express my sincerest gratitude and appreciation to those who have been generous to guide, advice, and support me for my PhD study through the valuable four years.

First of all, I would like to extend my thanks to both Dr. **Mark Wilkinson** and Prof. **Dave Fernig** for your wealth of knowledge and experience that help me tremendously with my work. They have been supporting me in terms of encouragement, guidance and advice, not only just in academic work, but also in challenging moment in my personal life.

I am grateful to my assessors, **Caroline Dart** and **Olga Mayans**, for being generous to guide me on the right track and offer the advice to enrich my study. Big thanks to **Ed Yates** and **Quentin Nunes** for helping with my experiments and editing my paper and thesis.

I would also like to thank the colleagues I have worked with in these four years, especially my friends from Lab D. I have had lots of advice from them, which has been very useful in my work. Their friendship has helped me have a better understanding of British culture and get used to British life, which made my PhD life very enjoyable. My special thanks to **Changye Sun**, for being my friend, which helped me get through many challenges over the past 4 years.

Last but not the least, special thanks to my parents, **Futian Li** and **Dexin Geng** for being so supportive of my decision to study aboard. It has been really hard for you to let me studying four years far from home and the financial support also increased your burden. I hope there is way to reward their faith in me.

Abstract

More than 883 extracellular proteins that bind heparan sulfate (HS) and heparin have been identified whose activities are regulated by their interactions with these polysaccharides. FGFs are heparin-binding proteins (the physiological ligand is HS) and the interactions of FGF with HS determine their transport between cells and the assembly of signalling complexes with their cognate receptor tyrosine kinases, FGFRs. The FGF family has expanded from two or three FGF ligands in the worm *C. elegans* and the fly *Drosophila* to 18 FGFs in vertebrates and mammals, which is directly linked to the more complex specifications required in the development for increasingly complex body parts. However, the level of specificity of the interaction of FGFs with HS is still debated. Previous work generally focused on just one or two FGFs and a limited repertoire of sugar structures, so a systematic investigation of the interaction of FGFs with heparin/HS is required to determine at what level, if any, there is specificity at the molecular level underlying these interactions. The strategy of this work was to use the evolutional relationship of the FGF family as a defined system to explore the specificity of interactions of FGFs and heparin/HS. Six FGFs (FGF3, FGF4, FGF6, FGF10, FGF17, and FGF20) from 4 subfamilies have been produced and purified as recombinant proteins, to investigate the interaction between FGFs and HS from two different perspectives. The polysaccharide structure required for binding to the FGFs was determined by differential scanning fluorimetry (DSF) using a library of chemically modified heparins and model glycosaminoglycans. The heparin binding sites on the FGFs were then identified by a lysine selective technique called ‘protect and label’. For systemic analysis of the interactions, all of the results obtained have been mapped on the FGF evolutionary tree deduced amino

acid sequence alignment, alongside previous work. This shows a clear pattern: FGF members from the same subfamily have a similar preference for binding particular subsets of HS/heparin structures and model glycosaminoglycans, and share similar secondary heparin binding sites on their surface. In contrast, FGFs from different subfamilies have a more divergent preference for binding structures in the polysaccharide and secondary binding sites on their surface. The secondary heparin binding sites (HBS) of FGF2 were mutated to begin the characterization of their functions. The properties of the mutants of FGF2's secondary HBSs (HBS2, HBS3) were measured in terms of their preference for binding structures in heparins, their ability to stimulate the phosphorylation of $p42/44^{MAPK}$ and cell proliferation. FGF2 (HBS2) mutant was found to be distinct to wild-type only in its interactions with low sulfated heparins where mutant FGF2 (HBS2) exhibited a stronger preference for N-sulfated heparin. For mutant FGF2 (HBS3), a larger sugar structure was required for binding than wild type FGF2. Finally, since both lysine and arginine residues in HBSs contribute to the interaction between protein and polysaccharide, an arginine targeted protect and label method was developed. Phenylglyoxal (PGO) was successfully used in the protection step and was demonstrated to be capable of achieving full arginine labelling. However, arginine labelling with 4-azidophenylglyoxal (APG) suffered from a ring expansion side reaction and this second step still needs to be optimized. Overall the thesis demonstrates that there is specificity in the interaction of FGFs and glycosaminoglycans. Although this is not a simple one-to-one code, it has clearly been subjected to the same natural selection that led to the expansion and diversification of the FGF family, the specificities of FGFs for particular isoforms of the FGFRs and selective activities of the FGF family in specifying the different structures and organs of the mammalian body.

Contents

List of Figures:

List of Tables:

List of Abbreviations

ACN: acetonitrile

APG: 4-azidophenylglyoxal

Arg: arginine

BF: Bradykinin fragment 2-9

BSA: bovine serum, albumin

CBB: Coomassie Brilliant Blue

CS: chondroitin sulfate

DMSO: dimethyl sulphoxide

DMEM: Dulbecco's modified Eagle medium

dp: degree of polymerization

DS: dermatan sulfate

DSF: differential scanning fluorimetry

DTT: dithiothreitol

ECM: extracellular matrix

EDTA: ethylenediamine tetra-acetic acid

FA: Fibronectin adhesion-promoting peptide

FCS: fetal calf serum

FGF: fibroblast growth factor

FGFR: fibroblast growth factor receptor

HA: hyaluronic acid

HBS: heparin binding site

HEPES: N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid

HRP: Horseradish peroxidase

HS: heparan sulfate

HSPGs: heparan sulfate proteoglycans

GAG: glycosaminoglycans

GlcNAc: N-acetylglucosamine

GlcNS: N-sulfoglucosamine

IdoA: L-iduronic acid

IPTG: isopropyl β-D-1-thiogalactopyranoside

LB: lysogeny broth

Lys-lysine

MTT: 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide

MW: molecular weight

NA: Neurotensin

NHS: N-hydroxysuccinimide

PAGE: polyacrylamide gel electrophoresis

PBS: phosphate-buffered saline

PCR: polymerase chain reaction

PDB: protein data bank

PGO: phenylglyoxal

PI: phosphatidylinositol

RE: reducing end

RM: routine medium

Rama: rat mammary

SD: standard derivation

SDM: step down medium

SDS: sodium dodecyl sulphate

SE: standard error

SRCD: synchrotron radiation circular dichroism

TB: Terrific Broth

TCA: trichloroacetic acid

TEMED: N,N,N’,N’,Tetramethylethylenediamine

TFA: trifluoroacetic acid

TM: Melting temperature

Tris: Tris (hydroxymethyl) methylamine

Tween 20: Polyoxyethylenesorbitan monolaurate

Chapter 1 General Introduction

1.1 The discovery of FGFs

In 1974, the term fibroblast growth factor (FGF) was first used by Gospodarowicz to describe a polypeptide found in pituitary and in brain that stimulated the initiation of DNA synthesis in resting 3T3 fibroblasts [1]. This was purified from both of these sources and was termed basic FGF (bFGF), because the protein was basic and stimulated mitogenesis in 3T3fibroblasts [2, 3]. Subsequently, brain tissue was found to also contain a second molecule with FGF activity, which was called acidic FGF (aFGF), because its isoelectric point was the opposite to that of bFGF's [4]. Thereafter, additional FGFs were identified and isolated by a variety of methods. According to the number-based nomenclature system, aFGF and bFGF were renamed as FGF1 and FGF2, respectively. FGF3 (int-2) [5], FGF4 (kFGF/hst) [6], FGF5 [7] and FGF6 [8] were originally isolated as oncogene products. FGF7 was identified as keratinocyte growth factor (KGF) with the distinctive properties of a mesenchyme-derived stimulator of epithelial cell growth [9]. FGF8 was isolated and characterized as causing unusual male sex hormone-dependent growth of mouse mammary carcinoma cells [10]. After a series of analyses of cDNA structure, Miyamoto *et al*. determined that glia-activating factor (GAF) was the ninth member of the FGF family, FGF9 [11]. FGF10 was isolated from rat embryos by homology-based polymerase chain reaction (PCR) [12]. Four further members of the FGF family, called FGF homologous factors (FHFs), FGF11 (FHF3), FGF12 (FHF1), FGF13 (FHF2) and FGF14 (FHF4) were found by a combination of random cDNA sequencing, data base searches, and degenerate PCR [13]. Following this, the Itoh

laboratory successfully isolated and identified FGF16, FGF17, FGF18, FGF20, and FGF21 by homology-based polymerase chain reaction [14], [15], [16], [17], [18]. FGF19, FGF22 and FGF23 were originally isolated by searching genome data bases [19], [20], [21]. Parallel discovery, allied to rapid analysis following completion of the human genome project resulted in the finding that human FGF19 correspond to mouse FGF15; there is no human FGF15 or mouse FGF19 [22].

1.2 Evolution of the FGF gene family

The *fgf* genes have not been found in unicellular organisms such as *Escherichia coli* and S*accharomyces cerevisiae*. After sequencing of the *Caenorhabditis elegans* and *Drosophila* genomes, two (*egl-17* and *let-756*) and three (*branch-less, pyramus and thisbe*) FGF genes were identified, respectively [23-27]. Six *fgf-like* genes were then identified in a basal chordate, the ascidian *Ciona intesinalis* that are considered to be potential ancestral genes of human FGF subfamilies, which include *fgf4-like, fgf5-like, fgf8-like, fgf9-like, fgf10-like* and *fgf13-like* [28]. Twenty-two *fgf* genes have been identified in the human, mouse and zebra fish *fgf* families [29]. The ancestral genes of *fgf* gene subfamilies have been hypothesized to be the result of gene duplication following the divergence of protostomes and deuterostomes [30]. A model of the evolutionary history of the *fgf* gene family has proposed that gene expansion presents in two phases (Fig. 1.1). In the first phase, there was gene duplication of the *fgf* gene from two or three to six genes during the early evolution of metazoans. In the second phase, the full complement of vertebrate genes was generated by two large-scale genome duplications [30, 31].

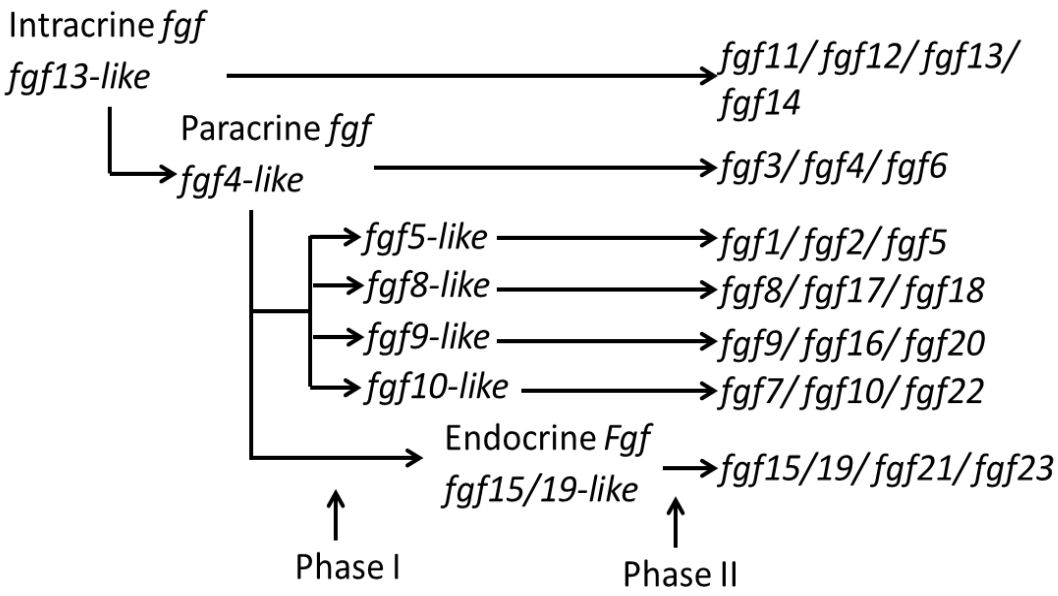


Figure 1.1 Possible evolutionary relationships between the *fgf* genes. The *fgf13-like* is the ancestral gene of the whole *fgf* gene family. The *fgf4-like* was generated by gene duplication from *fgf13-like* at an early stage. In the first phase (phase I), *fgf5-like*, *fgf8-like*, *fgf9-like*, *fgf10-like* and *fgf15/19-like* were generated by gene duplication. All of these expanded into three or four members by two large-scale gene duplications during phase II.

1.3 Phylogenetic and gene location analysis of the human/mouse *fgf* gene family

The *fgf* evolutionary relationship has been determined by the analysis of gene location on chromosomes [29, 30, 32, 33]. Seven *fgf* subfamilies have been suggested by gene location analysis: *fgf1, fgf2* and *fgf5*; *fgf3, fgf4* and *fgf6*; *fgf7, fgf10* and *fgf22*; *fgf8, fgf17* and *fgf18*; *fgf9, fgf16* and *fgf20*; *fgf11, fgf12, fgf13* and *fgf14*; *fgf15/19, fgf21* and *fgf2* (Fig 1.1).

Apart from gene location analysis, phylogenetic analysis has also been applied to determine the relationship of FGFs from their amino acid sequence (Fig 1.2) and also indicates that there are seven subfamilies: *fgf1* and *fgf2* (*fgf 1* subfamily); *fgf4, fgf5* and *fgf6* (*fgf4* subfamily); *fgf3, fgf7, fgf10* and *fgf22* (*fgf7* subfamily); *fgf8, fgf17* and *fgf18* (*fgf8* subfamily); *fgf9, fgf16* and *fgf20* (*fgf9* subfamily); *fgf11, fgf12, fgf13* and *fgf14* (*fgf11* subfamily); *fgf15/19, fgf21,* and *fgf23* (*fgf19* subfamily) (Fig. 1.2). The members of the *fgf8, fgf9, fgf11* and *fgf19* subfamilies are consistent with the phylogenetic and gene location analyses. However, *fgf5* and *fgf3* are indicated to be members of *fgf4* subfamily and *fgf7* subfamily, respectively, by phylogenetic analysis. The phylogenetic relationship based on sequence maps to functional similarities of the FGFs [34, 35] and it is in this context that FGF subfamilies will be discussed here.

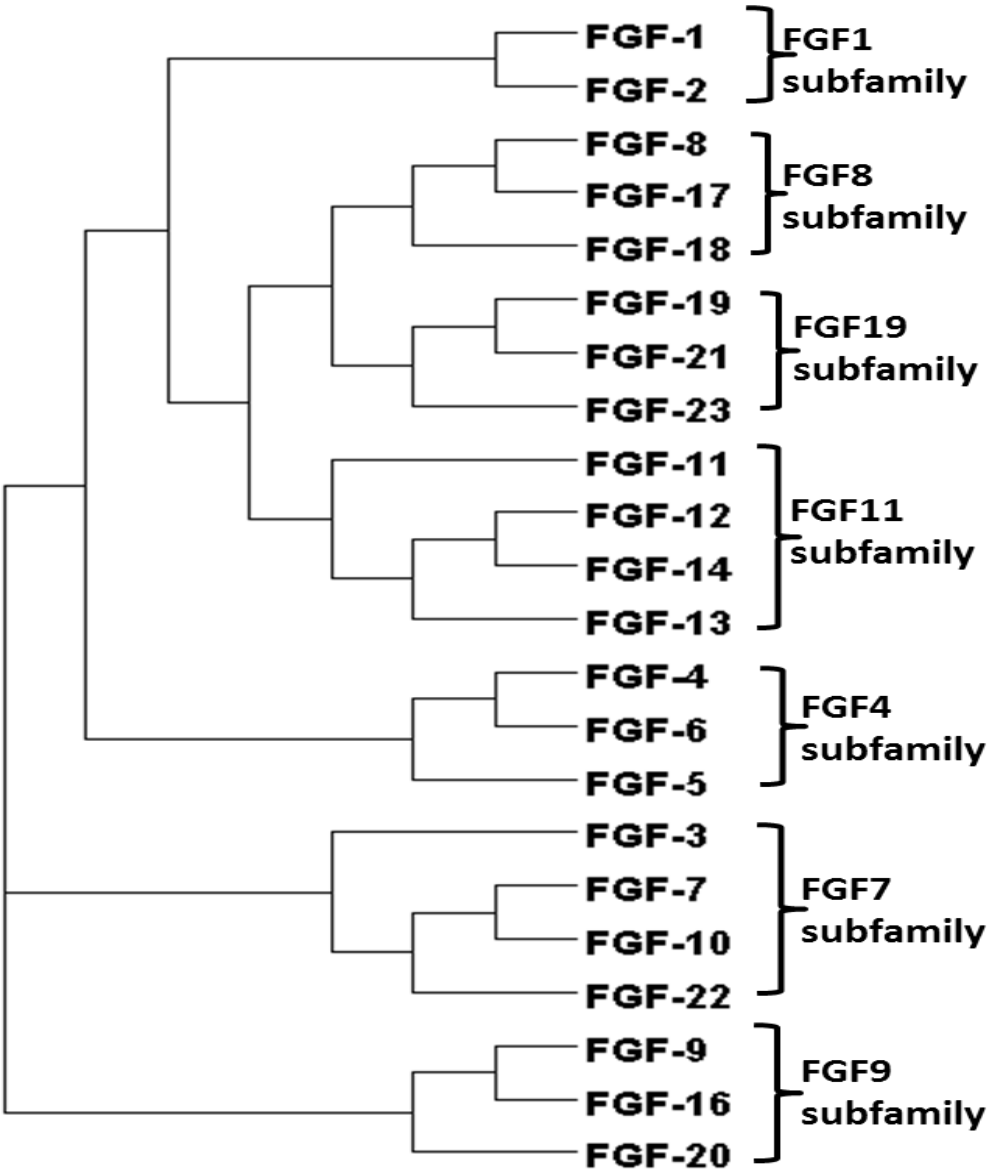


Figure 1.2 Evolutionary relationship tree of FGF family. According to amino acid sequence, Dendroscope was used to show that FGF family is divided into seven subfamilies. The branch lengths relate directly to the evolutionary relationship of FGFs.

1.4 FGF ligands

1.4.1 FGF ligand structure

The molecular weights of the FGFs range from 17 to 34 kDa in vertebrates, whereas it reaches 84 kDa in *Drosophila*. All FGFs share an internal core of similar structure with 28 highly conserved, and 6 invariant amino acid residues [36]. X-ray crystallography of FGFs shows that the FGF family possesses a similar folding pattern to the interleukins IL-1β and IL-1α [37]. FGFs possess a β trefoil structure, which is formed by three sets of four β strands connected by loops (Fig. 1.3) [38]. A variety of studies have demonstrated that the primary heparin binding site of FGF2 is formed by the strand β1/β2 loop, strands β10 / β11 loop, strand β11 and strands β11 / β12 loop (Fig. 1.4) [38-42]. The receptor tyrosine kinase (FGFR) binding site involves the strand β8-β9 loop and is distinct from the heparin binding site. This suggests that receptor binding and the binding of heparin and HS are likely to be physically separated [22, 31, 38].

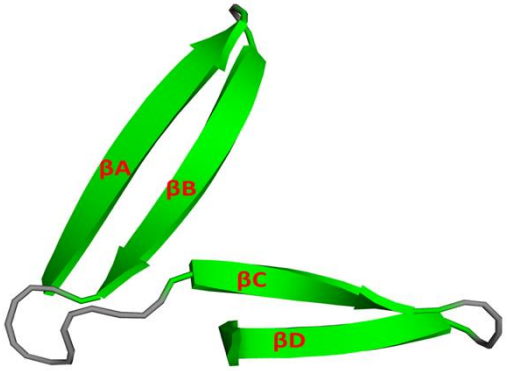


Figure 1.3 Schematic diagram of the core structure unit of the β-trefoil motif. The first ascending strand (βA) is connected to a descending strand (βB). The following "horizontal" strand (βC) is connected to the return strand (βD). Three of these units arranged around a pseudo three fold axis of symmetry form the β trefoil.

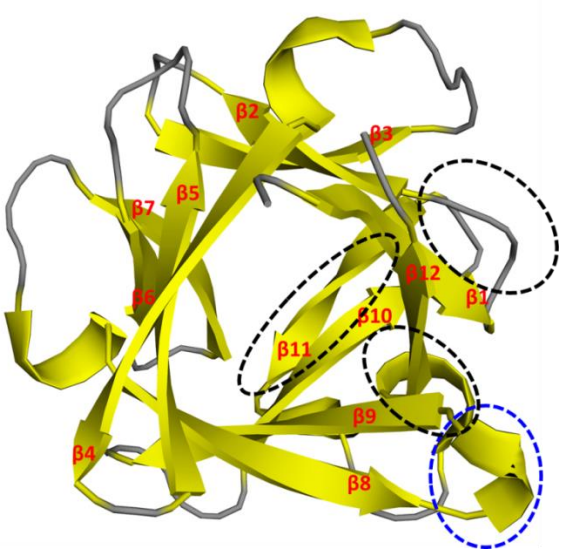


Figure 1.4 Ribbon diagram of crystal structure of FGF2. PDB ID: 2FGF [38]. The β trefoil structure is made up of 12 anti-parallel β strands. The structure of FGF2 was generated from the crystallographic coordinates with PyMol. The heparin binding sites are highlighted in black. The FGFR binding site is highlighted in blue.

1.4.2 The functions of FGFs

According to their mechanism of action, FGF activities can be classified into three types: intracrine, paracrine and endocrine. The paracrine FGF subfamilies are the FGF1 and FGF2; FGF3, FGF7, FGF10 and FGF22; FGF4, FGF5 and FGF6; FGF8, FGF17 and FGF18; FGF 9, FGF16 and FGF20. Most of these members of the FGF family have a cleavable N-terminal secretory signal peptide for export through the classical endoplasmic reticulum (ER)/ Golgi pathway. However, FGF1 and FGF2, which have no secretory signal peptide and have been shown to be exported from cells through mechanisms independent of the conventional endoplasmic reticulum - Golgi pathway [43-46]. Although members of FGF9 subfamily share the classical secretory pathway, they do not have a cleavable signal sequence. Instead, they possess a bipartite secretory signal sequence for their interaction with the secretory machinery [47, 48].

The paracrine FGFs all bind strongly to HS, a major glycosaminoglycan in pericellular and extracellular matrix, through their heparin binding sites. This interaction limits their radius of diffusion so consequently these FGFs exert their biological effects near their source of synthesis. The intracrine FGFs are the members of the FGF11 subfamily that are expressed in cells as intracellular proteins. The intracrine FGFs have been shown to have a crucial role in regulation of the excitability of granule neurons [49, 50]. The FGF19 subfamily members have weak or no binding to HS, so are able to perform their function in an endocrine manner [51]. They use α and β-Klotho as their co-receptor to bind to their cognate receptors [52-54]. Most *fgf* genes have been knocked out in mouse and the phenotypes indicate that in

this context they all have at least one non-redundant biological function (Table 1.1) [33, 55, 56].

Table 1.1 Summary of some physiological roles of FGF.

FGF1	Insulin resistance	Nutrient homeostasis
FGF2	Loss of vascular tone Slight loss of cortex neurons	Cardiovascular, skeletal and neuronal development
FGF3	Inner ear agenesis in humans	Inner ear development
FGF4	Embryonic lethal	Cardiac valve leaflet formation Limb development
FGF5	Abnormally long hair	Hair growth cycle regulation
FGF6	Defective muscle regeneration	Myogenesis
FGF7	Matted hair Reduced nephron branching in kidney	Branching morphogenesis
FGF8	Embryonic lethal	Brain, eye, ear and limb development
FGF9	Postnatal death Gender reversal Lung hypoplasia	Gonadal development Organogenesis
FGF10	Failed limb and lung development	Branching morphogenesis
FGF16	Embryonic lethal	Heart development
FGF17	Abnormal brain development	Cerebral and cerebellar development
FGF18	Delayed long-bone ossification	Bone development
FGF19	Increased bile acid pool	Bile acid homeostasis Lipolysis Gall bladder filling
FGF20	No knockout model	Neurotrophic factor
FGF21	No knockout model	Fasting response Glucose homeostasis Lipolysis and lipogenesis
FGF22	No knockout model	Presynaptic neural organizer
FGF23	Hyperphosphataemia Hypoglycaemia Immature sexual organs	Phosphate homeostasis Vitamin D homeostasis

Reference for each FGF (FGF1 [56]; FGF2 [57-62]; FGF3 [63]; FGF4 [64-66]; FGF5 [67-69]; FGF6 [70, 71]; FGF7 [72, 73]; FGF8 [74-76]; FGF9 [77, 78]; FGF10 [79]; FGF16 [80]; FGF17 [81]; FGF18 [82, 83]; FGF19 [84-88]; FGF20 [89]; FGF21 [90-94]; FGF22 [95]; FGF23 [96-101]).

1.5 Heparin/Heparan sulfate

1.5.1 Heparin/heparan sulfate structure

Heparan sulfate (HS) is a member of the family of glycosaminoglycans (GAGs), containing repeating disaccharide units joined by 1-4 linkages. As well as HS, the GAG family includes hyaluronic acid (HA), chondroitin sulfate (CS), dermatan sulfate (DS) and keratan sulfate (KS). The complex structure of HS distinguishes it from other GAGs [102]. The repeating units of HS consist of a glucuronic acid (β-D-GlcA) or its C5-epimer α-L-IdoA, and α-D-glucosamine (GlcN). The glucosamine may be N-acetylated (GlcNAc), N-sulfated (GlcNS), or unsubstituted (GlcN). O-sulfation can occur at position 2 of the uronic acid and position 3 and 6 of the glucosamine (Fig. 1.5). Heparin shares the same repeating disaccharide units as HS, but presents a far more homogeneous and higher level of sulfation [103]. Since it shares similarity in structure and is commercially available, heparin or its derivatives are often used as models for HS in studies of sugar-protein interactions [104].

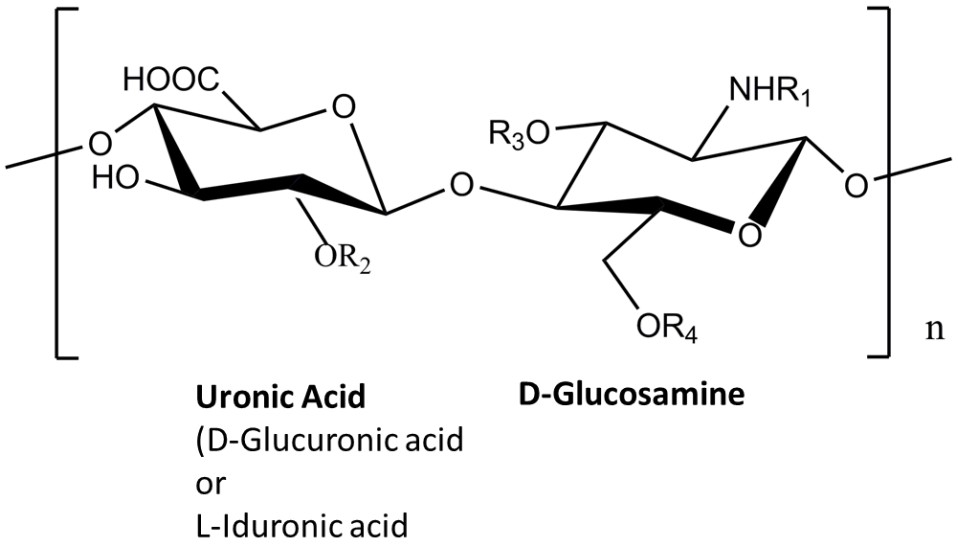


Figure 1.5 Heparin/HS disaccharide units. The repeating units of HS consist of a uronic acid (β-D-GlcA) or its C5-epimer α-L-IdoA, and D-glucosamine (α-GlcN). R_1 can be H, $COCH_3$ or SO_3^- ; $R_{2/3/4}$ can be H or SO_3^-.

1.5.2 Heparin/HS biosynthesis

The pathway of biosynthesis of HS can be divided into three sets of reactions: chain initiation, chain polymerization and polymer modification. Its biosynthesis occurs in the Golgi apparatus and begins with a linker tetrasaccharide being assembled on a specific serine residue of newly translated core proteins. Subsequently, the polymerization of the heparin/HS chain is performed by alternate addition of glucuronic acid and N-acetyl glucosamine to the non-reducing end of the growing polysaccharide. The repeated addition of these two saccharides is catalyzed by the tumor suppressor EXT family genes (EXT1 and EXT2) that form a stable heterodimeric complex [105, 106]. Afterwards, the sugar chain undergoes various modifications carried out by a series of enzymes. Firstly, the dual activity of the N-

deacetylase / N-sulfotransferases (NDSTs) catalyses the replacement of the N-acetyl group of glucosamine with a sulfate group. The presence of N-sulfate marks this disaccharide and the adjacent ones for further modifications. The C5 epimerization of glucuronic acid to iduronic acid through the C5 epimerase has been shown to only happen to the GlcA substrate that is attached to the reducing end of a GlcNS residue [107]. O-sulfation can then take place on position 2 of the uronic acid and position 3 and 6 of the glucosamine, catalysed by 2-O, 3-O and the 6-O sulfotransferases [108, 109]. Any GlcNS can be the substrate for 6-O sulfotransferases, whereas the 2-O-sulfotransferases cannot modify the IdoA attached to GlcNS(6S), which indicates that the 6-O-sulfation of GlcNS follows the 2-O-sulfation of IdoA [110]. The action of the N-deacetylase / N-sulfotransferases (NDSTs) is clustered. Thus, they will ignore a sequence of saccharide, and then act on alternating disaccharides, followed by every sequential disaccharide, and then alternating ones and so on. Since all the other modifications depend on the presence of N-sulfated glucosamine, the result is that HS chains have a domain structure: NA domains with no sulfation structure, NS domains of highly sulfated structures, and NA/NS domains comprising mixed regions of GlcNAc and GlcNS. The sulfated structures are of functional significance, forming the protein binding sites. Heparin chains contain more GlcNS residues than HS (the ratio of GlcNS to GlcNAc is around 4:1 in heparin and 1:1 in HS) and also contain more sulfated structures at an average of 2.4 sulfates/disaccharide than HS, which has O-sulfate structures in the range of 0.2-0.7 O-sulfate/disaccharide [103]. The sequential application of the biosynthetic enzymes does not readily explain such disaccharides as GlcA-GlcNAc6s or IdoA-GlcNAc6s, which have been found. An alternative view, implying enzymes capable of such conversion, has been proposed based on removal substrate specificities [111].

1.5.3 The functions of Heparin/HS

The HS chains are always synthesized attached to a core protein to form heparan sulfate proteoglycans (HSPGs) [112]. It is the core protein that directs the HS chains to specific locations, the cell surface or the extracellular matrix, where the polysaccharide chains bind and regulate the activity of proteins (883 HS/heparin-binding proteins identified to date [113, 114]). Thus, a large body of evidence demonstrates the function of HSPGs in many signalling pathways. For example, mutations in the *Drosophila* of genes encoding sugarless (*sgl*) and sulfateless (*sfl*) have been shown to have similar phenotypes as Wingless (Wg) or Hedgehog (Hh) signalling mutants [115]. In addition, the HS biosynthetic enzymes in *Drosophila*, Tout-velu (*ttv*), Sister of ttv (*sotv*) and brother of ttv (*botv*) are also required for Wg signalling gradient formation [116]. Indeed, studies of morphogen activity have shown the function of HSPGs in regulation of morphogen movement and intercellular trafficking [117]. A model of morphogen mobility suggests morphogen proteins can move from regions of high concentration to ones of low concentrations regions by transferal from one HSPG to the next in the pericellular matrix [118].

HS regulates several aspects of FGF function. The interaction of paracrine FGFs with HS is absolutely required for the formation of an efficient signalling complex with the FGFRs. Thus, HS was shown to be required for FGF2 to stimulate cell division in heparan sulfate-deficient Chinese hamster ovary (CHO) cells [119]. The same result was obtained by treating cells with heparitinase or culturing cells in sodium chlorate to block sulfation of heparan [120]. However, experiments on overexpression of branchless in Drosophila suggested that FGFR can be activated by high levels of FGF in the absence of HS and biochemical signalling by FGF2 has

been measured in cell cultures in the absence of HS [121-123]. Nonetheless, in all cases the usual paracrine FGF biological response (development of the trachea, cell division) required HS. Thus, the above work is also the first genetic evidence to show that HS is essential for normal FGF signalling *in vivo* [122]. Additional studies have shown that the function of HS may be to stabilize the active receptor complex for enough time for it to signal [36]. Moreover, structure defined oligosaccharides have also shown to differently activate FGFs, which indicates that different FGFs are likely to possess distinct affinities for unique HS structures [124, 125].

The role of heparan sulfate intracellular has also been explored. Intracellular heparan sulfate was first identified in the nucleus of hepatoma cells [126]. This HS was found to be enriched in 2-O sulfated glucuronic acid residues and to correlate with the control of cell growth and to originate from the extracellular space [126-128]. Since, intracellular HS binding proteins have been identified. For example, the catalytic activity of topoisomerase I, which changes the superhelical state of duplex DNA, has been shown to be inhibited by interaction with heparan sulfate [129]. Indeed, many nucleic acid binding proteins also bind heparin and presumably heparan sulfate. This is not surprising, since the sugar phosphate backbone of nucleic acids is analogous the sulfated glycosaminoglycans. Recently, the translocation of 3-O-sulfated heparan sulfate from the extracellular space to the neuronal cytoplasm has been shown, caused by the overexpression of 3OST2. This heparan sulfate binds to tau, which leads to the latter's hyperphosphorylation, a molecular change that drives Alzheimer's disease [130]. The intracellular presence of HS is established and some interactions demonstrated, such as that with tau. However, many of the interactions,

particularly with nucleic acid binding proteins have often only been demonstrated in vitro and whether they occur in the cell remains to be established.

Other functions of HS are demonstrated by genetic diseases associated with mutations in the genes encoding some of the biosynthetic enzymes. The EXT1 and EXT2 genes, whose products catalyse the elongation of the HS chain, have been shown to be tumour suppressor genes. Thus, studies of sporadic and exostosis derived chondrosarcomas show that a loss of heterozygosity for chromosomes 8 and 11 markers linked to EXT1 and EXT2 loci causes the disease [131, 132]. Idiopathic hypogonadotropic hypogonadism (IHH) is associated with mutations in heparan sulfate 6-O-sulfotransferase 1 (HS6O1) and has been linked to idiopathic hypogonadotropic hypogonadism pathogenesis by both Kallmann syndrome 1 (KAL1) -dependent and -independent mechanisms [133].

1.6 FGF receptors

1.6.1 FGF receptors

The FGFRs are encoded by five *fgfr* genes, *fgfr1-fgfr4* and *fgfr-L1*, which have been identified in humans, mice and zebrafish. Four of these receptors, *fgfr1-fgfr4* encode tyrosine kinases (*ca*.800 amino acids) [31, 55]. One and two *fgfrs* genes were identified in *C. elegans* (*egl-15*) and *Drosophila* (breathless and heartless), respectively. The receptor contains an extracellular region with three immunoglobulin domains (I, II, and III), a transmembrane helix, and a split intracellular tyrosine kinase domain (Fig.

1.6). The related *fgfr* gene, *fgfr-L1* (also called *FGFr5*) lacks an intracellular tyrosine kinase domain.

There are a large number of alternative RNA splicing events that generate different FGFR isoforms. The latter include two or three Ig loop isoforms that lack the D1 Ig domain and in some cases, e.g. *Drosophila heartless*, the gene only encodes this two Ig loop isoform [31]. In mammals there are three different exons encoding the C-terminal half of D3, Ig loop 3 in FGFR 1, 2 and 3. One, the "a" exon, encodes a secreted receptor extracellular domain while the "b" and "c" exons alter the sequence of D3 of the transmembrane receptor and provide a means to impact on ligand selectivity, such that the *fgfr* gene product can recognise preferentially different FGF ligand subfamilies [134, 135]. Immunoglobulin domain I and the adjacent acid box have no independent FGF ligand binding activity, but play an important role in receptor auto-inhibition [136] and in binding an alternative set of ligands, the cadherins [137]. FGF ligands mainly interact with the immunoglobulin domains II, III and the linker region between II and III of the receptor [31].

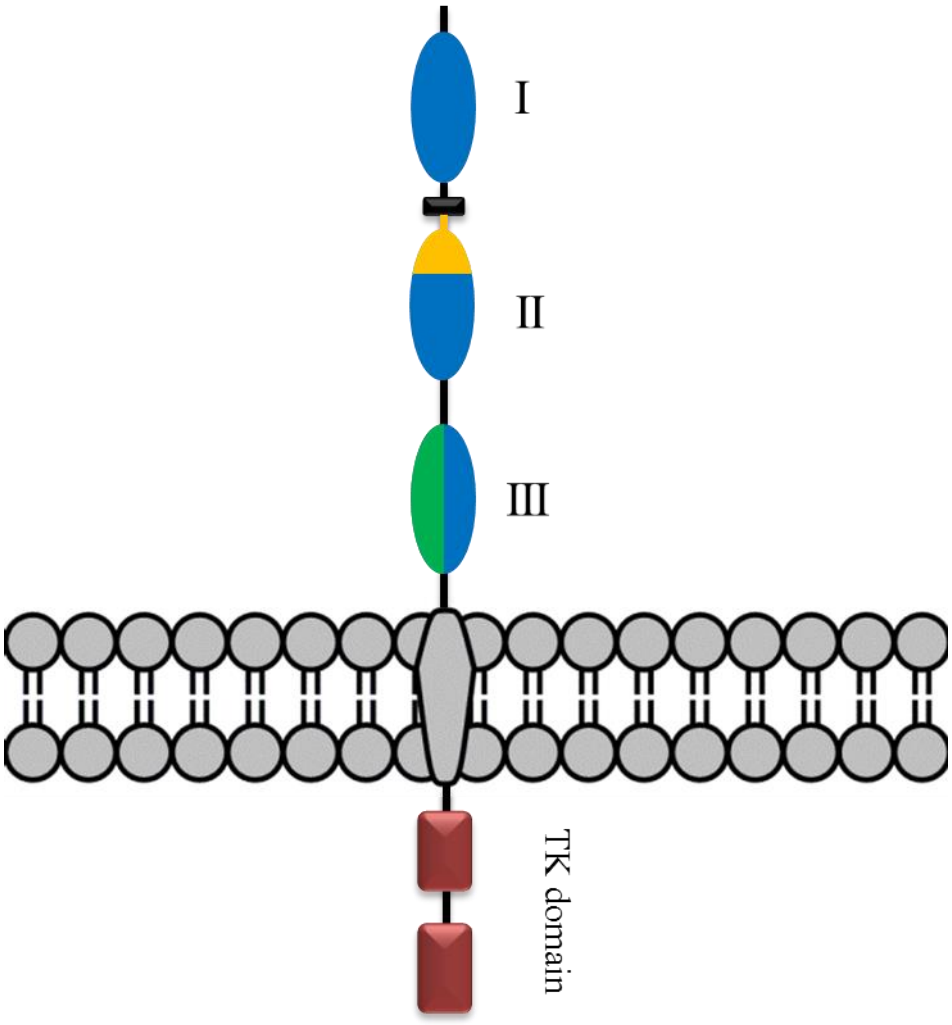


Figure 1.6 FGFR domain structure. There are three immunoglobulin domains (I, II, and III). The heparin binding site is colored orange on domain II. The black box between domains I and II is the acid box. FGF binds to Ig II and most FGFs ligands interact with the adjacent surface of Ig III. The green region is the alternatively spliced second half of immunoglobulin domains III. TK = tyrosine kinase.

1.6.2 FGF signalling pathway

The dimerisation of FGFR induced by engagement of the FGF ligand and HS co-receptor allows the transphosphorylation of the activation loop of the intracellular tyrosine kinase domain. Subsequent transphosphorylations by the activated FGFR kinase occur on other tyrosine residues on the intracellular domain of the receptor that act as docking sites for the initiators of intracellular signalling pathways. In addition the FGFR kinase phosphorylates tyrosine residues on the adaptor protein fibroblast growth factor receptor substrate 2α (FRS2α) and on the phosphatase SHP2. FRS2 binds the juxtamembrane region of the FGFR and this interaction does not require phosphorylated tyrosine, while SHP2 binds to phosphorylated FRS2. The growth factor receptor-bound protein 2 (GRB2) binds to either phosphorylated tyrosines on the FGFR or on FRS2. The adaptor protein GRB2-associated binding protein 1 (GAB1), commonly associated with the hepatocyte growth factor/scatter factor receptor MET also bind to FRS2, and when phosphorylated on tyrosine residues by the FGFR acts as a further binding site for GRB2. The bound GRB2 in turn binds son of sevenless (SOS), which recruits RAS, leading to activation of mitogen-activated protein kinases (MAPK) *via* the kinases RAF and mitogen activated protein kinase kinase (MEK). GAB1 is essential for binding and activating phosphatidyl inositol 3 kinase (PI3K), which causes the downstream phosphorylation and activation of the AKT pathway. In terms of the signal outcome of the FGFR, the GRB2-MAPK pathway is associated with cell division, whereas the PI3K/AKT pathway is associated with cell survival and cell fate determination [138-140]. Another well-established pathway is that of PLCγ, which binds directly to a phosphorylated tyrosine residue in the FGFR (tyr-766 in FGFR1). PLCγ is phosphorylated by FGFR kinase and then generates inositol triphosphate (IP3) and

diacyglycerol (DAG) by the hydrolysis of phosphatidylinositol, leading to the IP3-mediated release of calcium from the endoplasmic reticulum. Calcium ions and DAG together cooperate to activate protein kinase C (PKC). Even though the physiological role of this pathway is not well elucidated, it has been shown to be involved in cell adhesion [141].

The RAS-MAPK, PI3K and PLCγ pathways are the most studied and best described ones leading from the activated FGFR. These pathways cause changes in the activity of downstream kinases, phosphatases and proteins that result in the biological output of the FGF ligand. Generally, this is cell division or cell migration [140-143], but this may be a bias due to these being the most studied changes in cell behavior in the growth factor field.

1.6.3 Specificity of FGF and FGFR interactions

Except for FGF1, which interacts with all FGFRs and their isoforms, all other FGFs exhibit a greater or lesser degree of receptor and receptor isoform selectively (Table 1.2). Although there is usually no absolute specificity (a notable exception is FGF7, Table 1.2), it is the case that FGFs from the same subfamilies have a clear preference to bind particular FGFR isoforms, whereas FGFs from different subfamilies differ in this respect (Table 1.2). The expression of at least some FGFR isoforms is tissue specific: for example, FGFR b isoforms are generally expressed in the epithelial compartments, while the FGFR c isoforms are normally produced in the mesenchymal tissue compartment. A general theme is that the FGFs expressed in epithelial or mesenchymal tissues interact with FGFR produced in the opposite tissue [55].

Table 1.2 FGFR binding specificities of FGFs. Data from Ornitz *et al*. and Zhang *et al*. [35, 51].

FGF subfamilies	FGF	FGFR activity
FGF1 subfamilies	FGF1	All FGFRs
	FGF2	FGFR 1c, 3c > 2c, 1b, 4
FGF4 subfamily	FGF4	FGFR 1c, 2c > 3c, 4
	FGF5	
	FGF6	
FGF7 subfamily	FGF3	FGFR 2b > 1b
	FGF7	
	FGF10	
	FGF22	
FGF8 subfamily	FGF8	FGFR 3c > 4 > 2c > 1c >3b
	FGF17	
	FGF18	
FGF9 subfamily	FGF9	FGFR 3c > 2c > 1c, 3b >4
	FGF16	
	FGF20	
FGF19 subfamily	FGF19	FGFR 1c, 2c, 3c, 4 (weak activity)
	FGF21	
	FGF23	
FGF11 subfamily	FGF3	No known activity
	FGF7	
	FGF10	
	FGF22	

1.7 Models for complexes of FGFs and their receptors

There have been many studies aimed at establishing what the exact role of HS may be in the formation of signalling complexes with the FGFR. Although there is currently a consensus view on the role of HS, only some of the contradictory body of data has been resolved, a number of proposals exist to explain how the signalling complex is formed, and open questions remain.

Based on the idea that the FGF ligand has to cause dimerisation of the FGFR to activate the tyrosine kinase, there are data indicating that binding of the FGF to HS causes dimerisation of the former. A dimeric ligand would enable receptor dimerisation. Heparin induced oligomerisation of FGF1 [144] and *cis* dimers of FGF2 was demonstrated on a heparin octasaccharide, but not on shorter structures such as a disaccharide or a tetrasaccharide [145, 146]. This suggested that the specific side-by-side heparin-induced FGF2 dimer would be the minimal active structure required for its biological activity. In support of the dimersation model, an asymmetric crystallographic co-complex of 2:1:2 FGF1: heparin decasaccharide: FGFR complex has been proposed [147]. This 'asymmetric' model was formed by the dimerization of two FGF-FGFR complexes induced by heparin. Moreover, the decasaccharide interacts with both ligands in two 1:1 FGF1-FGFR2 complexes, but with just one of two receptors (Fig. 1.7B). In this and the symmetric model (see below) the atomic contacts between molecules are preserved, but the relative orientation of the receptors is quite different. Due to the lack of the contact between proteins, the role of heparin seems to be critical in the assembly of the 'asymmetric' complex [147]. Biophysical analysis was performed to validate the asymmetric model. This indicated a cooperative dimerisation of FGF1 on heparin derived oligosaccharides.

However, the dimerisation of FGFs on heparin and derived oligosaccharides requires a very high concentration of FGFs, where it is well established that the multivalency of the polysaccharide enables multiple FGFs to bind [148, 149].

An early biophysical analysis of FGF2 and FGFR1 using isothermal titration calorimetry, ultracentrifugation, molecular modeling and site-directed mutagenesis indicated that FGF2 binds to two receptors and one HS (FGF2: FGFR1: HS, 1:2:1) [150]. The affinity of FGF2 for FGFR1 was increased 10-fold by the presence of the heparin. In this model, the second FGFR1 was recruited to a secondary low affinity binding site on FGF2 with heparin stabilizing the complex (Fig. 1.7A). Since the second binding site of FGF2 has 250-fold lower affinity compared with the primary binding site mutation of the second binding site did not change measurable FGFR binding, but did cause a reduction of its mitogenic activity [151]. A similar effect was observed in FGF7 when point-mutations were proposed in the low affinity binding sites located loop between strands β9 and β10 [152]. This model, called "the growth hormone model" due to its mechanism being analogous to that of growth hormone, has since fallen out of favour, but there are as yet, no data to contradict it. One uncertainty, however, is whether some of the residues contributing to the secondary FGFR binding site on the FGF2 are also involved in the packing of the hydrophobic core of the protein, hence the absence of mitogenic activity seen when these residues are mutated, which may be due to a loss of stability of the FGF2.

The current consensus model is called the symmetric model and is based on data from co-crystals of a 2:2:2 complex of FGF2, a heparin-derived hexasaccharide and the extracellular Ig loops Ⅱ and Ⅲ of FGFR1, this is called the 'symmetric' model (Fig. 1.7C) (PDB ID 1FQ9) [112]. Primarily, the heparin oligosaccharide interacts with

FGF2 and FGFR1 to form a 1:1:1 complex. After that, dimerization of the two complexes is mediated by direct FGFR-FGFR interaction and secondary interaction between ligand in one ternary complex and receptor in the other ternary complex, with the FGF-FGFR interaction regulated by the sugar. It is important to note that these secondary contacts between FGF and FGFR are different from those in the growth hormone model. Since the non-reducing ends of both oligosaccharides face each other, this model was also called 'two-end' model [113]. The direct FGFR-FGFR and the secondary FGF-FGFR interaction were not sufficient for appreciable dimerization of the FGFR without the polysaccharide [112]. It is worth noting that, although the arrangements of the components differ in the symmetric and asymmetric models, the atomic contacts between FGF, heparin oligosaccharide and FGFR are very similar. Interestingly, in the symmetric model, the atomic contacts of the low affinity secondary FGFR binding site of FGF2 of the growth hormone model are free and exposed to solvent.

There may be more than one functional complex. For example, Zhu *et al*. (2010), provided evidence for different signalling complexes forming on fibroblasts with FGF2 and the endogenous FGFR1 [121]. These different intracellular signals are identified due to differences in signalling observed for complexes of FGF2: FGFR1 (no HS), FGF2: HS: FGFR1 and high FGF2 concentrations: HS: FGFR1.

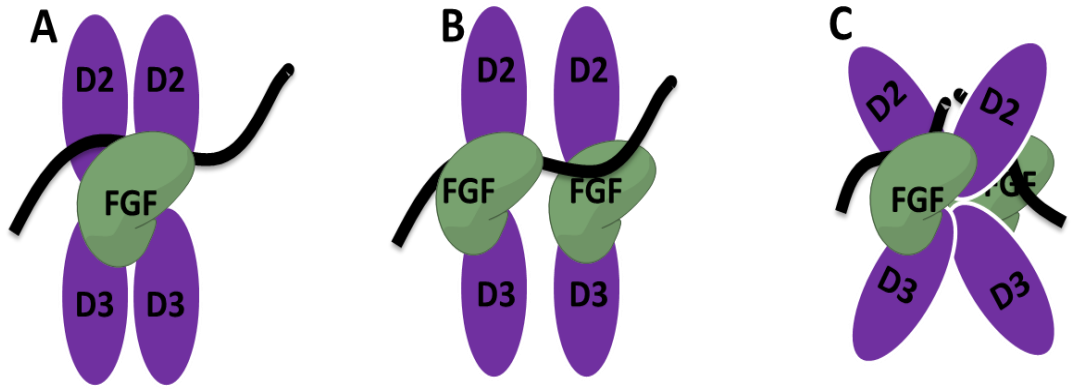


Figure 1.7 Models of FGF signalling complexes. Heparin chains are coloured black. FGF is coloured green. For FGFR, only the Ig II (D2) and Ig III (D3) domains are shown, coloured in purple. A, the growth hormone model; B, the 'asymmetric' model; C, the 'symmetric' model.

One fundamental difference between the 'symmetric' and 'asymmetric' model is the conformation of the invariant proline residue in the D2-D3 linker region. The invariant proline in the 'symmetric' model assumes a *trans* conformation, whereas the proline in the 'asymmetric' model is in *cis*. The *cis* and *trans* proline conformations were proposed to represent the active and inactive state of FGFR, respectively [147]. These two models have also been analysed in solution using size-exclusion chromatography, analytical ultracentrifugation and mass spectrometry. This work suggests that both structures were present, but that the 'asymmetric' model complex predominates [153]. However, the FGFR with *cis* proline was used, so an alternative interpretation, given that the symmetric complex was detected, is that the latter is physiologically relevant. Moreover, the disruption of the low affinity FGFR binding site of FGF caused a reduction in cell signalling, which supported the 'symmetric' model [154]. Analysis of the D2-D3 linker region in crystal structure 1E0O has shown that the *cis* conformation of proline is stabilized by a large excess

of Ni^{2+} ions in the crystallization buffer, so pushing the equilibrium between these two proline conformations towards to cis, which supports the idea that this may be an artefact of expression and purification [155].

1.8 Specificity of FGF and HS interactions

Heparan sulfate is a key regulator that controls the functions of a large number (883) of extracellular regulatory proteins [114, 156] (Section 1.5.3). FGFs are heparin-binding proteins and interactions with HS determine their transport between cells, as well as the assembly of signalling complexes with cognate FGFRs. In mammals, there are 18 FGFs involved in mediating cell-cell communication by interaction with HS and FGFR on target cells (Section 1.3). In the fly *Drosophila* and the worm *C. elegans*, there are just two or three FGF ligands and one *fgfr* alongside HS (Section 1.2). It seems reasonable that to develop and maintain the more complex bodies of vertebrates and mammals, having 18 FGFs and 5 *fgfr*s with HS is necessary and has arisen by natural selection. In studies on binding of modified heparin fragments (all containing N-sulfate groups) to FGF1, FGF2 and FGF4, the results showed that FGF2 required 2-O-sulfate rather than 6-O-sulfate for binding. FGF1 was more promiscuous and could bind structures with either 2-O-sulfate or 6-O-sulfate and FGF4 could bind at high concentration oligosaccharides lacking 2-O-sulfate or 6-O-sulfate [124, 125, 157]. Moreover, experiments are carried out *in vivo* have shown that FGFs clearly bind to different tissue compartments, e.g., in mammary gland FGF2 was found to bind more extracellular matrix HS surrounding duct and less to the matrix surrounding the terminal end buds where the ductal tree is elongating and branching [158]. Moreover, differences have been observed in the effects of HS on the

growth-stimulatory effects of FGF1 and FGF2, for example in 3T3 cells [159].

It remains the case, however, that the level of specificity for binding HS is still debated. One school, drawing on the perceived high specificity of the interaction of antithrombin III with a pentasaccharide in heparin, contends that there are rare and unusual sequences of saccharides in HS, and related heparin, responsible for high affinity binding of the proteins [160, 161]. Another school contends that largely non-specific ion-exchange interactions underpin protein binding by the polysaccharide [162, 163]. After crystallographic analysis of the amino acid residues located in the binding region of heparin and FGFs, it was shown that no single residue in the primary canonical heparin binding site is completely conserved throughout the FGF family [40], indicating that there has either been genetic drift or natural selection. In early studies, an octasaccharide from heparin had been suggested to be the minimal length oligosaccharide to induce the assembly of FGF2-FGFR1/2 and its biological activities [125, 164, 165], though later heparin-derived tetrasaccharides were found to be functional [123]. However, with oligosaccharides derived from HS, octasaccharides were the minimal active length [166]. In contrast, the deduced dimerization of FGFR2b induced by FGF7 required heparin-induced oligosaccharides longer than dp 8 [167]. Chemically modified heparins were used with FGF1 and FGF7 to show that FGF1-dependent dimerization of FGFR2b was enabled by 6-O- and 2-O-desulfated heparins, whereas neither of these polysaccharide could enable FGF1-FGFR1 or FGF7-FGFR2b to stimulate cell growth, which suggested that there may be a high degree of specificity in the formation of ligand-receptor complexes [167]. In the most comprehensive work, FGFs from five subfamilies were characterized in terms of their heparin binding specificity at different levels: binding parameters, the heparin structure required for binding, the identification of the heparin binding sites in the

FGFs and the secondary structure changes induced by heparin binding. With most of the subfamilies represented by just one member, these data led to the hypothesis that the differences in the preference of individual FGFs for particular heparin structures and the type of secondary heparin binding site in the FGFs may be reflected in the evolutionary relationship of the different subfamilies [34].

1.9 Aims

FGFs often serve as a model for understanding the consequences of the interaction of HS with other proteins. Thus, the question of the specificity between FGFs and HS has wide implications. Previous work has not been very systematic, generally involving just one or two FGFs and a limited repertoire of GAG structures [124, 125, 162, 168]. The FGF family of growth factors has expanded into clearly defined subfamilies through a series of genome duplications and these have clear functional relationships e.g. Table 1.1. Thus, the FGF family can be used as a defined system, subjected to natural selection, to determine the level of specificity of interactions of FGFs with GAGs. The existing work with six FGFs from five subfamilies suggests that there is specificity in FGF- heparin interactions and this reflects the evolution of the FGF family members, which parallels the specificity of FGF ligands for FGFRs [34, 169]. However, in this work only four of the five heparin binding FGF subfamilies was analysed and for three of these there was only one FGF represented. Thus, the idea that there is specificity in the interactions of FGFs with HS and that the polysaccharide binding properties of the FGFs have been subjected to the some selection pressure that led to their FGFR-binding specificities and functional specialization remains to be established.

A major aim of this thesis is to substantiate or otherwise the above hypothesis. To achieve this, at least two members of each paracrine subfamily have been produced to provide a far more comprehensive coverage of the FGF family. The specificity in their interactions of these FGFs with the polysaccharide has been determined in two ways. Firstly, the preferences of FGFs for particular sugar structures has been explored using differential scanning fluorimetry (DSF) and a library of chemically modified heparins, heparin derived oligosaccharides and model glycosaminoglycans. Secondly, the protect and label approach has been used to map primary and secondary heparin binding sites in the FGFs [170]. The secondary heparin binding sites (HBS2 and HBS3) of FGF2 have also been characterized by exploring the sugar structure preference and the biological activity of HBS mutants. Finally, a new arginine targeted protect and label technique has been developed, which will be able to identify these residues, which make a substantial contribution to the ionic bonding capacity of HBSs.

Chapter 2 General materials and methods

2.1 Electrophoresis

2.1.1 Agarose electrophoresis

Agarose gels (1.2 %, w/v) were made by melting agarose (0.4 g) (Bioline, London, UK) into 48 mL TAE buffer (40 mM Tris-Cl, 20 mM acetic acid, and 1 mM EDTA, pH 8). When the agarose had cooled slightly, 4 μL 10,000x SYBR (New England Biolab, UK) was added into the agarose. The molten agarose was poured into the gel making tray and allowed to cool. The samples and DNA ladder (1 kb New England Biolab, Herts, UK) were loaded on the gel after the gel tray was placed in the gel tank and covered with TAE buffer. Electrophoresis was carried out at 100 V, 30 min for each gel.

2.1.2 SDS-PAGE

Samples were analyzed by SDS-PAGE by mixing them with 1/5 volume of 5x SDS-PAGE loading buffer (50 % (v/v) glycerol, 10 % (w/v) SDS, 25 % (v/v) 2-mercaptoethanol in 0.3 mM Tris-Cl, pH 6.8 and coloured with bromophenol blue). Then, they were heated at 95°C for 3 min and centrifuged for 5 min at 12,000 x g. Typically, 10 μL sample and SDS-PAGE marker (SDS7-1VL, Sigma) were loaded onto a 12 % (w/v) SDS-polyacrylamide gel (Table 2.1) and the gels were run at 30 mA (per gel), 200 V for 50 min with running buffer (50 mM Tris-Cl, 192 mM glycine and 0.1 % (w/v) SDS).

Table 2.1 SDS-PAGE

Resolving Gel (ingredients for 12 % gel)	Total 10 mL
Acrylamide/ bis-acrylamide stock (30 %, w/v)	4.0 mL
Tris-HCl (3 M), pH 8.85	2.5 mL
Water	3.5 mL
10 % Sodium dodecyl sulphate (SDS) w/v	100 µL
TEMED (N, N, N', N', Tetramethylethylene diamine)	10 µL
Ammonium persulphate 50 mg/mL (freshly made)	100 µL
Stacking Gel	Total 10 mL
Acrylamide/ bis-acrylamide stock (30 %, w/v)	1.3 mL
Tris-Cl (1.25 M), pH 6.8	1 mL
Water	7.7 mL
10 % Sodium dodecyl sulphate (SDS) w/v	100 µL
TEMED	20 µL
Ammonium persulphate 50 mg/mL (freshly made)	100 µL

2.1.3 Western Blot

After separation of polypeptides on 12 % (w/v) SDS-polyacrylamide gels, the gels were packed against a PVDF membrane in a sandwich with pads. Proteins were transferred at 100 V for 60 min in transfer buffer (25 mM Tris-Cl, 192 mM glycine, 20 % (v/v) ethanol). The membrane was then blocked with 5 % (w/v) skimmed milk powder in TBST (50 mM Tris-Cl, 150 mM NaCl, 0.05 % (v/v) Tween 20, pH 7.5) for 1 h. The membrane was washed three times with TBST to remove excess milk proteins and incubated overnight with phospho-p$44/42^{MAPK}$ antibody (Cell signalling

Technology, Beverly, MA, USA) (1: 1000 dilution in TBST) on a shaker at 4°C. Following three washes with TBST, HRP-linked anti-rabbit IgG (Sigma, Gillingham, Dorset UK) (1: 5000 of a 1 mg/mL stock) was added to the membrane in TBST and incubated for 1 h at room temperature. After three washes with TBST, the membrane was covered with 1 mL ECL solution (GE Healthcare Life Sciences, Buckinghamshire, UK) and signal was detected with Hyperfilm (Fujifilm UK Ltd. Bedford, UK). The same membrane was re-blotted with anti-β-actin (Sigma, UK) as control after stripping with a solution containing 0.1 % (w/v) SDS, 1.5 % (w/v) glycine, and 1 % (v/v) Tween 20. The membrane was incubated twice with this solution at room temperature for 7 min, followed by washing twice with PBS (137 mmol/L NaCl, 2.7 mmol/L KCl, 10 mmol/L $Na_2HPO_4 \bullet 2\ H_2O$, 2.0 mmol/L KH_2PO_4, pH 7.4). After two more washes with TBST, the membrane was ready for further incubations with antibodies.

2.1.4 Coomassie Staining and Destaining

The gels were incubated in Coomassie stain (50 % (v/v) methanol v/v, 10 % (v/v) acetic acid, 0.25 % (w/v) CBB R-250) for 60 min, then soaked in de-staining buffer (30 % (v/v) methanol, 10 % (v/v) acidic acid) until the bands became clear.

2.1.5 Silver staining

After electrophoresis, the gels were soaked in fixative (40 % (v/v) ethanol, 10 % (v/v) acetic acid) for 1 hour. Then, the gels were rinsed twice in 10 % (v/v) ethanol for 5 min and washed in water three times for 5 min each. The gels were then incubated in 0.02 % (v/v) silver nitrate solution for 30 min. After rinsing in water for 5 s, the gels were washed once in freshly made developer solution (2.5 % (w/v) Na_2CO_3, 0.03 % (v/v) formaldehyde) until the solution went brown. The gels were transferred to fresh

developer buffer and left on a shaker until bands were stained to the required intensity. Stop solution (1 % acetic acid) was used to terminate staining and gels were then washed with water six times for 5 min each. The gels were rinsed with freshly made reducer (0.6 % sodium thiosulphate, 0.3 % potassium ferricyanide, 0.1 % sodium carbonate) until the background became clear. Then the gels were quickly washed in a large volume of water to remove the reducing buffer. Finally, the gels were washed five times for 5 min each with water.

2.2 cDNA cloning

2.2.1 Polymerase chain reaction (PCR)

PCR was used to amplify cDNAs in 50 µL reactions, the reaction mixture and PCR cycle settings are shown in Table 2.2.

Table 2.2 PCR reaction setting

A, PCR reaction mixtures

Mixture	Volume
Hot start polymerase buffer	5 µL
25 mM $MgCl_2$	3 µL
dNTPs	5 µL
PCR Grade Water	31 µL
Forward Primers	2 µL
Reverse Primers	2 µL
DNA template	1 µL
Hot start polymerases (New England Biolabs)	1 µL
Total	50 µL

B, PCR cycles

	Stage1	**Stage2**			**Stage3**	
Temperature (°C)	98	95	Tm	70	70	4
Time (s)	120	20	10	20	600	∞
Cycles	1	30	30	30	1	Finish

2.2.2 DNA digestion

The target DNA was digested by two corresponding enzymes at 37°C overnight. The ligation mixture is shown in Table 2.3.

Table 2.3 Digestion system

Mixture	volume
DNA	5 μL
NE Buffer	2 μL
Water	1 μL
NcoI-HF (New England Biolabs)	1 μL
BamHI-HF (New England Biolabs)	1 μL
Total	10 μL

2.2.3 Ligation

The insert and vector plasmid were ligated by T4 ligase (M0202L, New England Biolads, UK) at room temperature overnight. The ligation condition is shown in Table 2.4.

Table 2.4 Ligation system

Mixture	volume
Vector	3.5 µL
Insert	3.5 µL
T4 ligase buffer	2 µL
T4 ligase	1 µL
Total	10 µL

2.2.4 Mutagenesis

The target fragments were designed by site-mutation of sequence of FGF2 and bought from Life Technologies. Target fragments amplified by PCR with designed primers were digested with NcoI-HF and BamHI-HF to produce the fragment to be ligated with the vector, digested with the same enzymes. The ligated sample was transformed into DH5 alpha cells and then large-scale cultures were used for isolation of the target plasmid.

2.3 Protein expression

2.3.1 Materials

Lysogeny Broth (LB) culture medium (Merck, Watford, UK).

2xYT culture medium: 1.6 % (w/v) enzymatic digest of casein, 1.0 % (w/v) yeast extract, 0.5 % (w/v) NaCl, pH: 7.2.

Terrific Broth (TB) medium: 1.2 % (w/v) tryptone, 2.4 % (w/v) yeast extract, 0.4 % (v/v) glycerol, 0.17 M KH_2PO_4, 0.72 M K_2HPO_4.

2.3.2 Competent cell preparation

A single colony of selected bacteria (DH5 alpha/C41/BL21 (DE3)) was cultured in 40 mL of LB broth overnight at 37°C until the absorbance reached 0.3-0.5. The culture were immersed in ice for 10 min and transferred to 4 pre-chilled sterile centrifuge tubes and then centrifuged at 3800 x g for 10 min, 4°C. The pellets were placed on ice for 10 min and re-suspended in 10 mL pre-chilled 0.1 M $CaCl_2$. These samples were placed on ice for a further 20 min, and then centrifuged, as above. Finally, pellets were re-suspended in 2 mL 0.1 M $CaCl_2$. Competent cells were stored at -80°C.

2.3.3 Bacterial transformation

Seventy ng pETM-11 DNA and 70 µL competent C41 (DE3) or BL21 (DE3) plysS were thawed on ice. They were then mixed together and incubated on ice for 30 min. After a process of permeabilization induced by heat shock at 42°C for 1 min and incubation for 2 min on ice, cells were added to 1 mL LB and cultured at 37°C for 60 min. The cells were collected by centrifugation (3800 x g) for 5 min and re-suspended with 100 µL of LB and 10-25 µL of the cell suspension was plated into LB-antibiotic (Ampicillin/Kanamycin) plates. These plates were incubated at 37°C overnight.

2.3.4 Miniprep

A single colony from the plates was inoculated and cultured with 10 mL of LB broth-antibiotics at 37°C overnight (shaken at 240 rpm). Plasmids were purified

using a Qiagen miniprep kit 250 according to the manufacturer's instruction (Qiagen, Manchester, UK).

2.3.5 Sequencing

Plasmid DNAs (80-100 ng/µL) were sequenced by GATC sequence service (GATC, UK).

2.3.6 Bacterial culture

A single colony was inoculated and cultured with 10 mL of LB/ 2xYT broth-antibiotic at 37°C overnight (shaken at 240 rpm). Five ml of this culture was transferred to 500 mL of LB/ 2xYT broth with antibiotic and then cultured at 37°C with shaking (240 rpm) until the absorbance at 600 nm reached 0.5-0.9. At this point, the cultures were induced with IPTG (1 mM final concentration) and grown at 37°C (240 rpm) for 3 h, or overnight at 19.5°C (180 rpm). The bacteria were transferred to pre-cooled 1 L centrifuge tubes and centrifuged for 15 min at 4,150 x g. The cell pellets were re-suspended in PBS and transferred to 50 mL tubes. The cells were then centrifuged again for 10 min at 3,800 x g. The supernatants were discarded and pellets were stored at -80°C.

2.4 Protein purification

2.4.1 Cell breakage

Samples were thawed on ice with 5 volumes of lysis buffer (equilibration buffer for the subsequent chromatography step) and transferred to a 50 mL glass beaker for sonication. Cells were disrupted by six 30 s cycles of sonication, each for 1 min on

ice. Broken cells were centrifuged again in pre-chilled 50 ml centrifuge tubes at 38000 x g for 30 min. The supernatant was then used for the isolation of protein.

2.4.2 Chromatography

2.4.2.1 Heparin affinity chromatography

A column with 3 mL heparin resin (Affi-Gel Heparin, BioRad, UK) was equilibrated with buffer (50 mM Tris-Cl, 0.4 M NaCl, pH 7.2) at a flow rate of 1 mL/min. After cell breakage, the supernatant was filtered through a 0.45 µM filter and applied to the column. The column was then washed with washing buffer Ah (50 mM Tris-Cl, 0.4 M NaCl, pH 7.2) and washing buffer Bh (50 mM Tris-Cl, 0.6 M NaCl, pH 7.2). Elution was achieved by high ionic strength elution buffers Ch (50 mM Tris-Cl, 1.0 M NaCl, pH 7.2) followed by Dh (50 mM Tris-Cl, 1.5 M NaCl, pH 7.2). The column effluent was monitored at 280 nM with an Econo UV monitor (Bio-rad, Hertfordshire, UK).

2.4.2.2 Nickel affinity chromatography

A column of ProBond nickel-chelating resin (Invitrogen, Paisley, Scotland, U.K.) was equilibrated with equilibration buffer (50 mM Tris-Cl, 0.02 M imidazole, 0.3 M NaCl, pH 7.0). Then, the supernatant of disrupted cells was loaded onto this column and washed with washing buffer An (50 mM Tris-Cl, 0.02 M imidazole, 0.3 M NaCl, pH 7.2) and then washing buffer Bn (50 mM Tris-Cl, 0.1 M imidazole, 0.3 M NaCl, pH 7.2) at a flow rate of 1 ml/min. A higher concentration of imidazole was used in the elution buffers Cn (50 mM Tris-Cl, 0.25 M imidazole, 0.3 M NaCl, pH 7.2) and Dn (50 mM Tris-Cl, 1.0 M imidazole, pH 7.2). The column effluent was monitored at 280nm with a Bio-Rad (Hemel Hempstead, UK) Econo UV Monitor.

2.4.2.3 Cation-exchange chromatography

The fractions from heparin affinity chromatography were diluted 6 fold with buffer AI (50 mM Tris-Cl, pH 7.2) and applied to a cation-exchange column (1 mL HiTrap SP HP) set up on an AKTA system (GE Healthcare). Proteins were eluted with a 0-1 M NaCl gradient in 50 mM Tris-Cl, pH 7.2. The flow rate was 0.5 mL/min and elution was monitored at 280 nm.

2.5 Mammalian cell culture

2.5.1 Cell lines

Rama (rat mammary) 27 fibroblast cells

2.5.2 Tissue culture reagents

BSA (bovine serum, albumin) (Sigma-Aldrich); Trypsin/EDTA solution: 25 mL versene (Life Technologies) containing 2.5 % trypsin (in PBS, w/v); PBS (phosphate buffered saline): 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.5 Mm KH_2PO_4, pH 7.2; RM (Routine medium): DMEM (Dulbecco's modified Eagle's medium) 1 × (Gibco Life Technologies), 10 % (v/v) FBS (foetal bovine serum), 0.75 % sodium bicarbonate (Invitrogen, Paisley, UK), 20 mM L-glucosamine (Gibco), 1.00 IU/mL, penicillin (Gibco), 100 μg/mL streptomycin (Gibco), 50 ng/mL insulin (Sigma-Aldrich), 50 ng/mL hydrocortisome (Sigma-Approach); SDM medium (step down medium): DMEM, 0.75 % sodium bicarbonate, 20 mM L-glucosamine (Gibco), 1.00 IU/mL, penicillin (Gibco), 100 μg/mL streptomycin (Gibco), 250 μg/mL BSA (Sigma, catalogue number 7030); Freezing medium: 7.5 % (v/v) DMSO (Dimethyl

sulphoxide) and 20 % (v/v) FBS in DMEM; Tissue culture Petri dishes (Nunc, Roskilde, Denmark).

2.5.3 Cell culture

Rama (rat mammary) 27 fibroblast cells were thawed at 37°C and transferred to a 25 mL universal tube. Twenty mL DMEM (with 20 % (v/v) FBS) was added slowly into the universal tube and then the cells were centrifuged at 800 rpm for 5 min. The supernatant was removed and the pellet was re-suspended in 10 mL pre-warmed RM and all of cells were then placed into culture dishes and incubated at 37°C.

The cells were cultured with RM at 37°C in 9 cm diameter dishes in a humidified atmosphere of 10 % carbon dioxide and 90 % air (both v/v). The cells were sub-cultured at a ratio of 1:8 when the density of cells reached 70 % confluence. Cells were washed with PBS twice and then detached by incubating in 1 ml Versene at 37°C for 1 min. Seven mL RM was applied to re-suspended detached cells and then 1 mL was transferred to new culture dish. To retain their growth characteristics, the passage range of cells used was between 30 and 40.

2.5.4 Cell counting

Detached cells (0.5 mL) were diluted to 20 mL with Coulter Isoton II diluent (Beckman Coulter). The cells were counted in a Coulter Electronics particle counter. Cells from each dish were counted three times.

2.5.5 Freezing cells

Detached cells were transferred to a universal tube and centrifuged at 120 x g for 5 min. After removing the supernatant, the cell pellet was resuspended with 1 mL RM medium and then stored in 1.5 mL cryotube at -80°C.

2.6 Identification of proteins using peptide mass fingerprinting

Gel slices from SDS-PAGE were washed with 50 % acetonitrile, 0.2 M ammonium bicarbonate pH 8.9 and then dried in a rotary evaporator. The slices were re-swollen in rehydration buffer (0.2 M ammonium bicarbonate pH 7.8, 2 M urea) containing trypsin and incubated at 37°C overnight. Peptides were extracted from the gel slices with 60 % (v/v) acetonitrile, 0.1 % (w/v) TFA, de-salted using Millipore C18 ZipTips (Merck, Watford, UK). The latter were pre-wetted with 100 % (v/v) acetonitrile (ACN) and then pre-equilibrated with 0.1 % (w/v) TFA. They were loaded on the Zip Tip and then washed with 10 µL 0.1 % (W/v) TFA. Finally, these peptides were eluted with 2 aliquots of 4-6 µL 50 % (v/v) CAN. The peptides were then subjected to MALDI-TOF (Waters Corp., Manchester, UK) analysis using hydrocinnamic acid as the matrix.

2.7 Size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS)

Analysis of the solution molecular mass were performed by separation of proteins on Superose 200 HR10/300 columns (GE Healthcare) connected in series with a Wyatt Dawn8^{+} and Wyatt Optilab T-rEX (Wyatt Technology) at 22°C. Samples (100 µL) were filtered and then applied to the column, which was developed in 150 mM NaCl buffer with either 50 mM HEPES or 50 mM Tris-Cl, both pH 7.4 at a flow rate of 0.75 mL/min.

2.8 Differential Scanning Fluorimetry (DSF)

The sugar structure required for binding to FGFs was measured by DSF. It was performed with a 7500 Fast Real Time PCR (RT-PCR) system (software version 1.4.0 Applied Biosystems, Life Technologies, Paisley, UK). The melting temperatures (Tm) of a variety of FGFs (at 5 µM) were determined in the presence of various sugars. Heparin (0-500 µM), heparin-derived oligosaccharides of degree of polymerisation (dp) 2 to 12, chemically modified derivatives D1-D9 (Fig 4 B), porcine mucosal heparin sulphate, HA, CS, DS, and cation modified heparins were used as 10-fold concentrated stock solution (1.75 mg/ml) in HPLC grade water. First of all, protein stock solution (3.5 µL), heparin derivatives (3.5 µL), phosphate-buffered saline (24.5 µL), and freshly prepared 100-fold stock solution Sypro Orange dye (3.5 µL) were added to a Fast Optical 96 Well Reaction plate (Applied Biosystems) maintained on ice. After sealing with Optical Adhesive Film (Applied Biosystems), the plate was directly analysed in the RT-PCR system. The heating cycle was performed as a gradient between 32 and 81°C in 99 steps of 20 s. First derivatives of the melting curves were calculated with Origin 7 (OriginLab Corp., Northampton, UK). For each heparin derivative, at least two experiments each with triplicate wells were performed and analyzed. The mean Tm and the standard error (S.E.) were calculated based on the six repeats. The formula used to normalized data was: Tm x-Tm PBS /Tm hep- Tm PBS, where Tm x is the Tm of protein in the presence of the heparin derivative; Tm PBS (or Tm Tris) is the Tm of protein itself, and Tm hep is the mean Tm of the protein in the presence of heparin. The relative stability of protein in PBS or Tris-Cl buffer was set to 0, while the relative stability of heparin was set to 1.

2.9 Identification of heparin binding sites by Protect and Label

2.9.1 Protection labelling on heparin

A heparin minicolumn was made by placing a plastic air filter at the end of P10 tip (P50) into which 30 µL AF—heparin beads (Tosoh Biosciences GmbH, Stuttgart, Germany; binding capacity of 4 mg ATIII/mL) was packed. A 5 mL syringe was used to pack the minicolumn and dispense buffer. The heparin column was equilibrated 4 times with 50 µL of PB 150 buffer (17.9 mM Na_2HPO_4, 2.1 mM NaH_2PO_4, 150 mM NaCl, pH 7.8). A minimum of 40 µg FGF protein was loaded onto the heparin column and the loading was repeated 3 times with the same sample. After binding, the column was washed with PB 150 buffer 4 times. To acetylate exposed lysines, the minicolumn was then quickly rinsed with 20 µL of PB 150 containing 50 mM sulfo-NHS-acetate and then incubated for 5 min with 20 µL of PB 150 containing 50 mM fresh sulfo-NHS-acetate at room temperature. After acetylation, the minicolumn was washed with 50 µL of PB 150 buffer and acetylated protein was eluted with 2 x 20 µL elution buffer (45 mM Na_2HPO_4, 5 mM NaH_2PO_4, 2 M NaCl, pH 7.8).

2.9.2 HBS Lysine Biotinylation

Acetylated protein was diluted with 200 µL of PB buffer (17.9 mM Na_2HPO_4, 2.1 mM NaH_2PO_4, pH 7.8) and concentrated with a 5 kDa-MWCO centrifugal filter (Sartorius Ltd., Epsom, UK) about 15 min at 9,000 x g. The volume was adjusted to 37.2 µL with PB buffer and the biotinylation was performed by the addition of 2.8 µL of 145 mM NHS-biotin in DMSO and a 30 min incubation at room temperature.

The biotinylation reaction was quenched with 4 µL of 1 M Tris-Cl, pH7.5. Then, the reaction was transferred to a desalting centrifugal column and covered with 70 µL HPLC grade water. After 2 minutes centrifugation at 1,200 x g, samples were stored at -80°C for 10 min and dried by centrifugal evaporation.

2.9.3 Protein Digestion

Dried sample was dissolved in 25 µL of 8 M urea, 400 mM NH_4HCO_3, pH 7.8 and 2.5 µL of 45 mM DTT and incubated for 15 min at 56°C. Then, the samples were carbamidomethylated in the dark with 2.5 µL of freshly made 0.1 M iodoacetamide for 15 min at room temperature. Proteins were diluted with 70 µL of HPLC grade water and digested overnight with 1µg of chymotrypsin.

2.9.4 Biotinylated Peptide Purification

A minicolumn was made by placing a plastic air filter at the end of P10 tip (P50). Then, forty µL Strep-Tactin Sepharose beads (IBA-Solutions for Life Sciences, Life Technologies, Paisley, UK) were packed in this column. It was then rinsed 4 times with 50 µL 500 mM urea, 25 mM NH_4HCO_3. Digested protein was diluted to a final volume of 400 µL with HPLC grade water. The sample was then loaded onto the minicolumn and re-cycled through three times to ensure efficient binding. The column was washed 3 times with 50 µL 500 mM urea, 25 mM NH_4HCO_3 and then washed twice with 50 µL HPLC water to remove unbound peptides. Biotinylated peptides were eluted with 20 µL 80 % (v/v) ACN, 20 % (v/v) TFA and 5 mM biotin.

2.9.5 Identification of Labelled Peptides

Biotinylated/acetylated peptides were made up to 0.5 % (w/v) TFA and desalted using C18 Zip Tips as in section 2.7. The samples were concentrated by rotary evaporation and analyzed by MALDI-TOF mass spectrometry.

Chapter 3 Expression, purification and characterisation of recombinant FGFs

3.1 Subcloning of FGF cDNAs

3.1.1 Materials

cDNA-cDNA encoding human FGF-4 (UniProt accession number P08620; residues 31-206) and FGF2 mutants HBS2 (208K/A, 219K/A, 249R/A, 251R/A, 252K/A) and HBS3 (160K/A, 163K/A, 164R/A) (all from Life Technologies) were inserted into vector pETM-11 (a kind of gift from Dr Paul Elliott, University of Liverpool, UK). pET-14b-FGF-2, pETM-11-FGF-3, pETM-11-FGF-7, pETM-11-FGF-16 and pET-14b(halo)-FGF-17, were prepared as described [171].

Table 3.1 Primers designed for amplifying FGFs.

FGFs	Primer sequence
FGF4:	Forward: 5'-TTCAGGGCGCCATGGCG-3'
	Reverse: 5'-GAGTGCGGCCGCGAA-3'
FGF2 (HBS2):	Forward: 5'-ATGGGTCATCATCATCACCATCATC-3'
	Reverse: 5'-GAGTGCGGCCGCGAATTCGGATCCT-3'
FGF2 (HBS3):	Forward: 5'-ATGGGCCATCATCACCATCACCATC-3'
	Reverse: 5'-GAGTGCGGCCGCGAATTCGGATCCT-3'

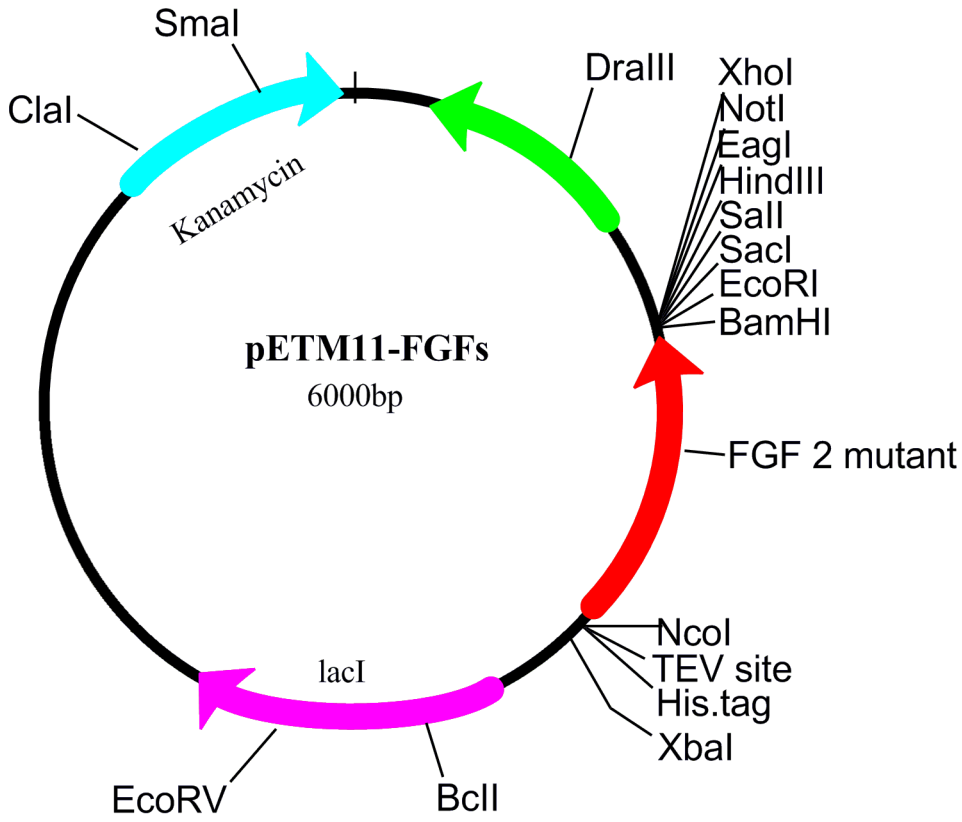
3.1.2 Methods

3.1.2.1 Production of FGF4 cDNA and the FGF-2 mutant cDNA

A cDNA encoding FGF4 with two flanking cleavage sites (NcoI, CCATGG and BamHI, GGATCC) was designed and expression was optimized by Life Technologies through reduction of mRNA secondary structure and adapting the DNA sequence to *E. coli* codon usage as far as possible. Two FGF2 mutants were similarly designed and optimized lysine and arginine residues of heparin binding sites were replaced with alanine. Depending on which heparin binding site was mutated, the mutants were named as FGF2 (HBS2) and FGF2 (HBS3).

3.1.2.2 Ligation

The cDNA fragments of FGF4 and of the FGF2 mutants were digested with NcoI-HF and BAMHI-HF to produce the insert which was then ligated with digested pETM-11 vector (Fig. 3.1, Section 2.22 and 2.2.3). The ligation product was transformed into DH5 alpha. All ligation products were sequenced to ensure they were identical to the designed DNA sequences.



Figures 3.1 Expression of plasmid map of FGF4 and FGF2 mutants. FGF4 (534 bp) and FGF2 mutants (546 bp) (Red) were ligated into pETM-11, which contain a 6xhistag and a TEV-site.

3.2 Paper: HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors

Changye Sun, Yong Li, Sarah E Taylor, Xianqing Mao, Mark C. Wilkinson, David G. Fernig. (2015). HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. PeerJ 3:e1060; DOI 10.7717/peerj.1060.

Contributions

Yong Li: The cloning of cDNA of Histag-FGF3, Histag-FGF7, Halotag/Histag-FGF16 and Halotag/Histag-17. Expression and Production of FGFs. Assisted with biological activities assay. Data analysis. Co-wrote the paper.

This paper has shown the process of DNA cloning, protein expression and purification of Histag-FGF3, Histag-FGF7, Halotag/Histag-FGF16 and Halotag/Histag-17. The following sections are supplementary data showing the process on FGF2, His-FGF2 mutants and His-FGF4.

HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors

Changye Sun[1,3], Yong Li[1,3], Sarah E. Taylor[1], Xianqing Mao[2], Mark C. Wilkinson[1] and David G. Fernig[1]

[1] Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool, UK
[2] Department of Oncology, Laboratory of Cellular and Molecular Oncology, Luxembourg Institute of Health, Luxembourg
[3] These authors contributed equally to this work.

ABSTRACT

The production of recombinant proteins such as the fibroblast growth factors (FGFs) is the key to establishing their function in cell communication. The production of recombinant FGFs in *E. coli* is limited, however, due to expression and solubility problems. HaloTag has been used as a fusion protein to introduce a genetically-encoded means for chemical conjugation of probes. We have expressed 11 FGF proteins with an N-terminal HaloTag, followed by a tobacco etch virus (TEV) protease cleavage site to allow release of the FGF protein. These were purified by heparin-affinity chromatography, and in some instances by further ion-exchange chromatography. It was found that HaloTag did not adversely affect the expression of FGF1 and FGF10, both of which expressed well as soluble proteins. The N-terminal HaloTag fusion was found to enhance the expression and yield of FGF2, FGF3 and FGF7. Moreover, whereas FGF6, FGF8, FGF16, FGF17, FGF20 and FGF22 were only expressed as insoluble proteins, their N-terminal HaloTag fusion counterparts (Halo-FGFs) were soluble, and could be successfully purified. However, cleavage of Halo-FGF6, -FGF8 and -FGF22 with TEV resulted in aggregation of the FGF protein. Measurement of phosphorylation of p42/44 mitogen-activated protein kinase and of cell growth demonstrated that the HaloTag fusion proteins were biologically active. Thus, HaloTag provides a means to enhance the expression of soluble recombinant proteins, in addition to providing a chemical genetics route for covalent tagging of proteins.

Subjects Biochemistry
Keywords Fibroblast growth factor, Recombinant protein expression, HaloTag, Fusion protein

Submitted 23 December 2014
Accepted 8 June 2015
Published 25 June 2015

Corresponding author
David G. Fernig, dgfernig@liv.ac.uk

Academic editor
Sandhya Visweswariah

Additional Information and Declarations can be found on page 21

DOI 10.7717/peerj.1060

INTRODUCTION

Of the 18 receptor-binding fibroblast growth factors (FGF), 15 also bind a heparan sulfate co-receptor and are classed as growth factors and morphogens. These are grouped into 5 subfamilies based on their protein sequence similarity (*Itoh, 2007*; *Ornitz, 2000*), and they regulate a myriad of processes in development, homeostasis and in some diseases (*Beenken & Mohammadi, 2009*; *Turner & Grose, 2010*). Recombinant FGFs provide a key tool to study their structure–function relationships, and labelling FGFs for microscopy

How to cite this article Sun et al. (2015), HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. **PeerJ 3:e1060; DOI 10.7717/peerj.1060**

has been important in probing the mechanisms of, for example, their transport (*Duchesne et al., 2012*; *Lin, 2004*; *Yu et al., 2009*). Chemical labelling has disadvantages compared to genetically encoded labelling, since with the latter it is easier to predict the structural and hence functional consequences of labelling, which can be achieved both *in vitro* and *in vivo*. While fluorescent proteins remain a mainstay of genetic labelling, they have limitations. These have been overcome, for example, by non-covalent tagging of proteins on hexahistidine sequences with Tris-Ni^{2+} nitriloacetic acid (*Huang et al., 2009*; *Lata et al., 2005*; *Tinazli et al., 2005*), which has allowed diverse labelling strategies, ranging from fluorescent dyes (*Uchinomiya et al., 2009*) and quantum dots (*Roullier et al., 2009*; *Susumu et al., 2010*) to gold nanoparticles (*Duchesne et al., 2008*). However, non-covalent coupling is reversible and exchange may occur in this instance with histidine-rich patches on endogenous proteins.

HaloTag is a mutant of a bacterial haloalkane dehalogenase, which reacts with chloroalkane ligands to form a covalent bond that represents the covalent intermediate of the enzyme's normal catalytic cycle (*Los et al., 2008*). Fluorescent dyes (*Los et al., 2008*) and quantum dots (*Zhang et al., 2006b*) carrying a chloroalkane group have been used to label HaloTag fusion proteins for fluorescence imaging. This approach is particularly versatile, since it combines the power of a genetically encoded tag (the HaloTag protein) with covalent labelling.

Consequently, we set out to produce N-terminal HaloTag fusions of different FGFs. In the course of this work, we observed that the N-terminal HaloTag fusion had a substantial effect on the expression of the more recalcitrant FGFs, consistent with the observation that HaloTag is a potential solubilisation tag for recombinant proteins (*Ohana et al., 2009*). Thus, whereas expression of FGF1 and FGF10 was somewhat reduced and that of FGF2 increased, expression of FGF7, which can be toxic (*Ron et al., 1993*), was no longer so, while expression of soluble FGF3, FGF6, FGF7, FGF8, FGF16, FGF17, FGF20 and FGF22 was markedly enhanced. This is in contrast to previous reports where FGFs such as FGF6 (*Pizette et al., 1991*), FGF8 (*Loo & Salmivirta, 2002*; *Macarthur et al., 1995*; *Vogel, Rodriguez & IzpisuaBelmonte, 1996*), FGF16 (*Danilenko et al., 1999*) and FGF20 (*Jeffers et al., 2002*; *Kalinina et al., 2009*) have been found to be mainly expressed in inclusion bodies, even as truncated proteins, and so require refolding. Thus, HaloTag provides not just a means to label proteins covalently and specifically, but is also a useful solublisation partner for the production of recombinant proteins.

MATERIALS AND METHODS

Materials

pET-14b vectors containing cDNAs encoding FGF1 and FGF2 and pET-M11 vector containing FGF7 cDNA were as described (*Xu et al., 2012*); cDNAs encoding FGF3, FGF10, FGF16, FGF17 and FGF20 were purchased from Eurofins Genomics (Ebersberg, Germany); cDNA encoding FGF6, FGF8 and FGF22 were purchased from Life Technologies (Paisley, UK); cDNAs encoding HaloTag was acquired from Kazusa DNA Research Institute (Kisarazu, Japan); Primers for PCR were from Life Technologies (Paisley, UK). All

Table 1 Peptide sequences of FGFs, the N-terminal HisTag constructs and the N-terminal HaloTag constructs. FGF names, sequences and amino acid numbering are according to the UniProt entry. FGF1 is an N-terminal truncated protein (*Ke et al., 1990*). FGF2 does not possess a secretory signal sequence, whereas there is no signal peptide recognised in Uniprot for FGF16 and FGF20; consequently full length protein sequence was expressed. For all other FGFs, the protein expressed was without the Uniprot determined secretory signal sequence. FGFx refers any one of the FGFs. TEV cleavage sites are in red.

Name	UniProt accession number	Residues in mature protein
FGF1	P05230	16–155
FGF2	P09038-2	1–155
FGF3	P11487	18–239
FGF6	P10767	38–208
FGF7	P21781	32–194
FGF8b	P55075-3	23–215
FGF10	O15520	38–208
FGF16	O43320	1–207
FGF17	O60258-1	23–216
FGF20	Q9NP95	1–211
FGF22	Q9HCT0	23–170
HisTag terminus (pET-M11)		MKHHHHHHPMSDYDIPTTENLYFQGA-[FGFx]
HaloTag and TEV site to conjoin with FGF sequence		MPEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYV WRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEAL GLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPE FARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREP FLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWG TPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLS TLEISGEPTTEDLYFQS-[FGFx]

of the protein sequences corresponding to the above cDNAs are listed in Table 1. Enzymes for cloning were from: NcoI, BamHI and T4 ligase (NEB, Hitchin, UK); KOD Hot Start DNA polymerase (Merck, Hertfordshire, UK); In-Fusion® HD Cloning Kit (Clontech, Takara Bio Europe SAS, Saint-Germain-en-Laye, France). Bacterial cells: DH5α, BL21 (DE3) pLysS and SoluBL21 were a gift from Olga Mayans, University of Liverpool. The sources of other materials were as follows: LB broth and LB agar (Merck, Hertfordshire, Germany); Soniprep 150 Plus (MSE, London, UK); Affi-Gel® Heparin Gel (Bio-Rad, Hertfordshire, UK), CM Sepharose Fast Flow, DEAE Sepharose Fast Flow, HiTrap Q HP column; empty disposable PD-10 Columns; ÄKTApurifier 100 plus (GE Healthcare, Buckinghamshire, UK).

For cell culture the following materials were used: Dulbecco's modified Eagle's medium (DMEM, Life Technologies), fetal calf serum (FCS, Labtech International Ltd, East Sussex, UK), 7.5% (w/v) sodium bicarbonate (Invitrogen, Paisley, UK), 200 mM L-glutamine (Gibco), 5 µg/mL insulin (Sigma-Aldrich, Dorset, UK), 5 µg/mL hydrocortisone (Sigma-Aldrich), bovine serum albumin (BSA; A7030, Sigma-Aldrich) for cell culture, cell culture dishes (Corning, Nottingham, UK). For SDS-PAGE and Western blotting: dried skimmed milk (Marvel, Spalding, UK), BSA (Fisher Scientific, Lougborough, UK), protease inhibitor (Roche, Burgess Hill, UK), phospho-p44/42 MAPK (T202/Y204) antibody (Cell Signalling, NEB, Hitchin, UK), monoclonal β-actin antibody

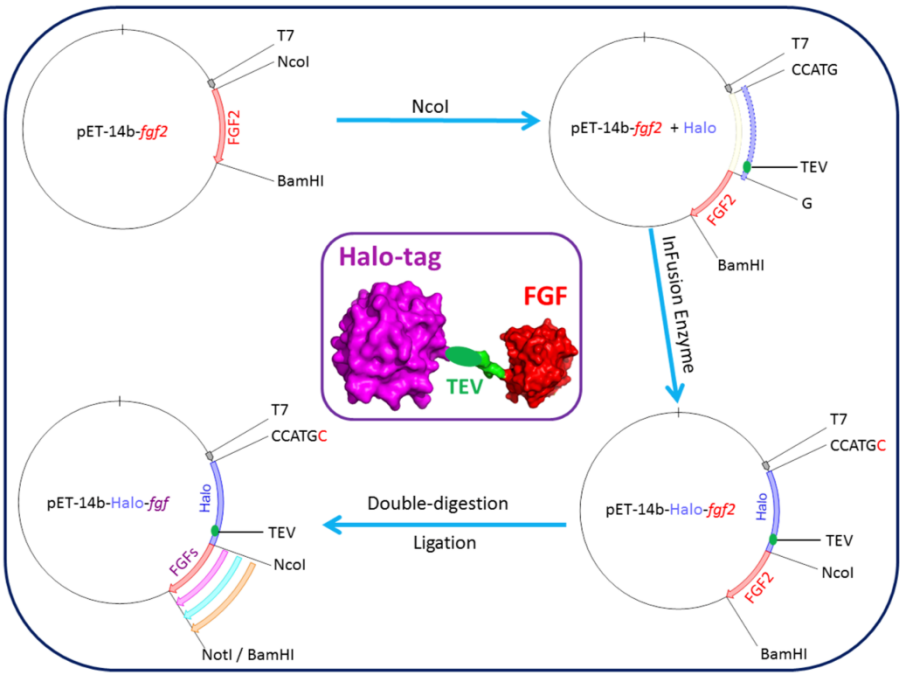


Figure 1 Cloning strategy for plasmids encoding Halo-FGFs. DNA encoding HaloTag was inserted 5′ of the FGF2 coding sequence with the In-Fusion HD enzyme. Subsequently, a NotI cleavage site was added 5′ to the BamHI site and other FGFs were exchanged into the plasmid using the digestion-ligation cloning method. A cartoon structure of Halo-FGF is presented in the middle of this figure.

(Sigma-Aldrich, Dorset, UK), anti-mouse IgG, horseradish peroxidase-linked antibody (Cell Signalling, NEB), polyvinylidene fluoride (PVDF) transfer membrane (Millipore UK, Hertfordshire, UK), enhanced chemiluminescence (ECL) Western blotting reagents (GE Healthcare, Little Chalfont, UK), Hyperfilm (GE Healthcare, Little Chalfont, UK).

DNA cloning of hexahistidine tagged FGFs (His-FGFs) and HaloTag tagged FGFs (Halo-FGFs)

DNA encoding FGF1, FGF3, FGF6, FGF8, FGF10, FGF16, FGF17, FGF20 and FGF22 was cloned into pET-M11 such that the resulting protein would have a N-terminal 6xhis tag followed by a tobacco etch virus (TEV) cleavage site (ENLYFQ). FGF2 and FGF7 DNA sequences were previously cloned into pET-14b and pET-M11, respectively (*Xu et al., 2012*).

A plasmid encoding Halo-FGF2 was produced by adding a HaloTag encoding DNA sequence in-frame 5′ to a DNA sequence encoding full-length FGF2. This construct was then used to produce the other DNAs encoding Halo-FGFs (Fig. 1). The plasmid pET-14b-*fgf*2 contains NcoI and BamHI cleavage sites 5′ and 3′ of *fgf*2, respectively. This vector was linearized by digestion with NcoI. The DNA encoding HaloTag (Fig. 1: blue insert) was amplified by PCR using the Halo-FGF2-Forward, AAGGAGATATA CCATGCCAGAAATCGGTACTG, and Halo-FGF2-Reverse,

Table 2 Concentrations of NaCl in 50 mM Tris-Cl buffer (pH 7.4) used for heparin affinity chromatography of FGFs. [NaCl] for lysate is the concentration of NaCl in the sample applied to the column.

Name	[NaCl] for lysate (M)	[NaCl] for wash (M)	[NaCl] for elution (M)
FGF1	0.6	0.6	1.0
FGF2	0.6	0.6	1.5
FGF3	0.3	0.6	1.0
FGF6	0.3	0.4	1.0
FGF7	0.3	0.3	1.0
FGF8	0.6	0.6	1.5
FGF10	0.6	0.6	1.0
FGF16	0.3	0.4	1.0
FGF17	0.6	0.6	1.0
FGF20	0.3	0.4	1.0
FGF22	0.6	0.8	1.5

TCCCGGCTGCCATGGAGCTCTGAAAGTACAGATC, primers (NcoI cleavage site underlined), and inserted into the linearized vector using In-Fusion enzyme. A TEV cleavage site (Fig. 1: green ellipsoid) was also included at the C-terminus of HaloTag to allow release of the FGF. A NotI cleavage site was also inserted 5′ of the BamHI to provide an additional 3′ cleavage sites for cloning. The other cDNAs (FGF1, FGF3, FGF6, FGF7, FGF8, FGF10, FGF16, FGF17, FGF20 and FGF22) were exchanged into the established pET-14b-*Halo-fgf2* plasmid by double-digestion with NcoI and BamHI/NotI enzymes and ligation using T4 ligase (Fig. 1).

Protein expression and purification of His-FGFs and Halo-FGFs

His-FGF7, because it is toxic like native FGF7 (*Ron et al., 1993*), was transformed into BL21 (DE3) pLysS (F– ompT hsdSB(rB–, mB–) gal dcm (DE3) pLysS (CamR)) for subsequent protein expression and purification. FGF2, the other His-FGFs and Halo-FGFs were transformed into SoluBL21(F– *omp*T *hsd*SB(rB–, mB–) *gal dcm* (DE3)). The bacteria containing FGF encoding plasmids were cultured at 37 °C until the OD600 values were between 0.4 and 0.6, and then protein expression at 16 °C was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacteria were harvested by centrifugation at 4 °C, 14,000 g for 10 min and the pellets frozen at −80 °C.

The bacterial pellets were resuspended with the corresponding 50 mM Tris-Cl lysate buffers (pH 7.4) (Table 2), and the cells were disrupted by 5–6 cycles of sonication (30 s sonication, 60 s pause) on ice. Cell debris and insoluble proteins were removed by centrifugation at 4 °C, 30,000 g for 30 min. Then, the presence of soluble FGFs was tested by analysis of whole cells, the supernatant and pellet by separation of polypeptides on 12% (w/v) SDS-PAGE and coomassie staining.

FGF2 and His-FGF7 were purified as described before (*Xu et al., 2012*). Soluble FGF1, FGF2, FGF3, FGF10, FGF16 and FGF17, including His-FGFs and Halo-FGFs, were loaded onto a 3 mL and the other soluble FGFs were loaded onto an 8 mL column of heparin

agarose. For each FGF, different concentrations of NaCl (in 50 mM Tris-Cl pH 7.4) were used for washing and elution (Table 2) by following the previous measurements on the electrolyte sensitivity of their heparin binding assessed by Western blot (*Asada et al., 2009*). The yields of His-FGFs and Halo-FGFs were quantified by measuring the absorbance at 280 nm and the level of impurities were estimated by analysis of coomassie stained SDS-PAGE gels with ImageJ-Analyze-Gels (*Ferreira & Rasband, 2012*). The soluble His-FGFs eluted from heparin affinity chromatography were further purified by Ni^{2+} affinity chromatography. Due to the negative charge on the surface of HaloTag and positive charge on the surface of FGFs, Halo-FGFs could bind to both cation- and anion-exchange stationary phases. Thus, Halo-FGF1, Halo-FGF2, Halo-FGF3, Halo-FGF7 and Halo-FGF10 were purified by chromatography on a 5 mL HiTrap Q HP column. Samples were applied in 0.15 M NaCl in PB buffer (2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and eluted with a gradient running to 0.8 M NaCl in the same buffer. Halo-FGF6 and Halo-FGF20 were purified by chromatography on a 3 mL column of CM Sepharose Fast Flow followed by a 3 mL column of DEAE Sepharose Fast Flow. Samples were again applied in 0.15 M NaCl in PB buffer and eluted with 0.4 M NaCl in the same buffer. The purified His-FGFs and Halo-FGFs were analysed by 12% (w/v) SDS-PAGE followed by coomassie staining.

Purification of FGFs by removing HaloTag from Halo-FGFs

To test the accessibility of the TEV cleavage site, some Halo-FGFs, including Halo-FGF2, Halo-FGF17, Halo-FGF6, Halo-FGF8 and Halo-FGF22 eluted with high concentration of NaCl in 50 mM Tris buffer from heparin agarose chromatography and Halo-FGF20 purified with heparin, DEAE and CM chromatography, were incubated with 2.5% (mol/mol) TEV protease at 4 °C overnight. In cases where the digestion products were cloudy, they were clarified by centrifugation for 30 min at 13,000 g, 4 °C. Samples were then analysed on a 12% (w/v) SDS-PAGE. The supernatants of the TEV digestions of Halo-FGF6 and of Halo-FGF20 were applied onto a 2 mL heparin agarose column, and washed as before (Table 2). FGF6 and FGF20 were eluted with PB buffer containing 1 M NaCl or 0.1 M arginine and 1 M NaCl, respectively. After TEV digestion, FGF17 was further purified on a 1 mL HiTrap SP HP cation-exchange column by washing with 0.3 M NaCl and eluting with 1 M NaCl, both in 50 mM Tris-Cl, pH 7.4. All of the fractions from the purification steps were analysed by 12% (w/v) SDS-PAGE.

Cell culture

Rama 27 cells were cultured in DMEM medium containing 10% (v/v) FCS, 4 mM L-glutamine, 0.75% (w/v) sodium bicarbonate, 50 ng/mL insulin and 50 ng/mL hydrocortisone (*Rudland, Twiston Davies & Tsao, 1984*). HaCaT cells were cultured in the same medium, but without insulin and hydrocortisone (*Boukamp et al., 1988*). Cell number was measured with a Z1 coulter particle counter (Beckman Coulter, High Wycombe, UK).

Measurement of p44/42^{MAPK} phosphorylation

Cells were cultured in 3 cm dishes until near confluence. Then, the dishes were washed twice with phosphate-buffered saline (PBS) and 2.5 mL step-down medium (SDM: DMEM with 250 ng BSA, 0.75% (w/v) sodium bicarbonate and 4 mM L-glutamine) was added for 24 h (Rama 27 cells) or 48 h (HaCaT cells). Rama 27 and HaCaT cells were then incubated with different FGFs for 15 min, as described in the figure legends. After the incubation, the cells were washed twice with ice-cold PBS and collected by scraping in 2X SDS-PAGE lysis buffer (4% (w/v) SDS, 20% (v/v) glycerol, 12% (v/v) Tris-Cl (pH 6.8), 2.5% (v/v) β-mercaptoethanol, 0.02% (w/v) bromophenol blue, 1 tablet of protease inhibitor and 6.8 mL distilled water). The cell lysates were heated for 10 min at 98 °C prior to SDS-PAGE.

Western blot

After separation by 10% (w/v) SDS-PAGE, polypeptides were transferred onto a PVDF membrane. The membrane was blocked with 5% (w/v) skimmed milk in 1X TBST (50 mM Tris-Cl, 150 mM NaCl and 0.05% Tween-20 (v/v), pH 7.5) for 2 h. After two washes with TBST, the membrane was incubated with phospho-p44/42^{MAPK} antibody (1:1,000 dilution in TBST) on a shaker overnight at 4 °C. Secondary anti-mouse antibody (1:1,000 dilution) was added to the membrane after three washes with TBST, 5 min each, for 1 h at room temperature. Following three washes with TBST to remove the excess secondary antibody, the membrane was covered with 1 mL ECL solution and signal was detected with Hyperfilm. The same membrane was stripped with 2.5% (w/v) SDS in TBST at 50 °C for 1 h and reblocked as above, before probing with β-actin antibody (1:10,000 dilution). The Western blot band intensity was quantified in the same way as SDS-PAGE bands and the signal intensities of phospho-p44/42^{MAPK} were normalised by dividing by the intensity of the band corresponding to β-actin and then by that of the BSA control samples.

Cell growth assay

Cell growth in Rama 27 fibroblasts was measured as before (*Smith, Winslow & Rudland, 1984*). Rama 27 cells were dispensed into a 24 well cell culture plates at 2,000 cells/well. After 24 h the cells were washed twice with PBS and cultured in SDM, as described for the p44/42^{MAPK} phosphorylation assay for 24 h. The SDM was then replaced and the appropriate proteins added, as described in the figure legend. After 68 h incubation, cells were trypsinised and the number of cells counted.

RESULTS AND DISCUSSION

Expression of soluble FGFs

Based on their relative expression and solubility properties, the FGFs were split into three different groups: FGFs that expressed well as soluble proteins (Group 1: FGF1, FGF2 and FGF10), FGFs that expressed at a low level (Group 2: FGF3 and FGF7), and FGFs that were insoluble when expressed in *E. coli* (Group 3: FGF6, FGF8, FGF16, FGF17, FGF20 and FGF22).

Group 1: soluble FGFs

After induction, bands corresponding to the expected molecular size of His-FGF1, FGF2 and His-FGF10 were apparent in the whole cell lysates (Figs. 2A, 2C and 2E, lane L, green arrow). His-FGF1 and His-FGF10 were expressed at a higher level than FGF2 in *E. coli* SoluBL21. After centrifugation of the cell lysates, bands corresponding to the molecular size of all three FGFs were mainly recovered in the soluble fraction (supernatant), rather than in the insoluble fraction (pellet; Figs. 2A, 2C and 2E, lanes S and P). Chromatography of the supernatants on heparin demonstrated that little of the expressed protein was present in the flow-through fraction (Figs. 2A, 2C and 2E, lane T). Weak bands corresponding to His-FGF1 and His-FGF10, but not FGF2, were observed in the wash fraction (Figs 2A and 2E, lane Wa), which may represent aggregated or less well-folded protein. The majority of the three FGFs was recovered in the high NaCl eluate (Figs. 2A, 2C and 2E, lane Hep), which demonstrated that these soluble FGFs bound heparin strongly. This indicated that they were likely to be properly folded, because the canonical, highest affinity heparin binding site of FGFs depends on the tertiary structure of the proteins (*Xu et al., 2012*).

The bands corresponding to Halo-FGF1, Halo-FGF2 and Halo-FGF10 were clearly observed in the whole cell lysates and these proteins were all highly expressed in SoluBL21 cells (Figs. 2B, 2D and 2F, lane L, red arrow). Similarly to the His-FGF1, FGF2 and His-FGF10, after centrifugation of the whole cell lysates, the bands corresponding to the three Halo-FGFs were observed in the soluble fractions (Figs. 2B, 2D and 2F, lanes S and P). Chromatography of the soluble fractions on heparin indicated that most of Halo-FGF2 and Halo-FGF10 had bound to the column, but there was a substantial amount of Halo-FGF1 in the flow-through (Figs. 2B, 2D and 2F, lane T). This may be due to the capacity of the column for Halo-FGF1 being lower than for His-FGF1. All three Halo-FGFs were eluted from the heparin affinity column at the expected NaCl concentration (Figs. 2B, 2D and 2F, lane Hep).

The yield of Halo-FGF1 and Halo-FGF10 was similar to that of the corresponding his-tagged proteins (Table 3). However, since the Halo-FGF proteins are considerably larger than the corresponding His-tagged FGF1 and FGF10, this represents a decrease in the molar amounts of FGF produced. In contrast, the yield of Halo-FGF2 was 4-fold higher (Table 3), which is only partly accounted for by the increased size of the fusion protein. The low yield of full-length FGF2 has been ascribed to the presence of secondary structure at the 5′ end of the FGF2 mRNA (*Knoerzer et al., 1989*), and the presence of the upstream HaloTag sequence may mitigate this effect.

Group 2: low expression proteins

The expression of His-FGF3 was weak, as was that of His-FGF7 (expressed in BL21 DE3 pLysS) due to its toxicity (*Ron et al., 1993*) (Figs. 3A and 3C, lane L, S and P, green arrow). Heparin chromatography of the supernatants demonstrated that the yields of soluble His-FGF3 and His-FGF7 were quite low (Figs. 3A and 3C, lane Hep; Table 3).

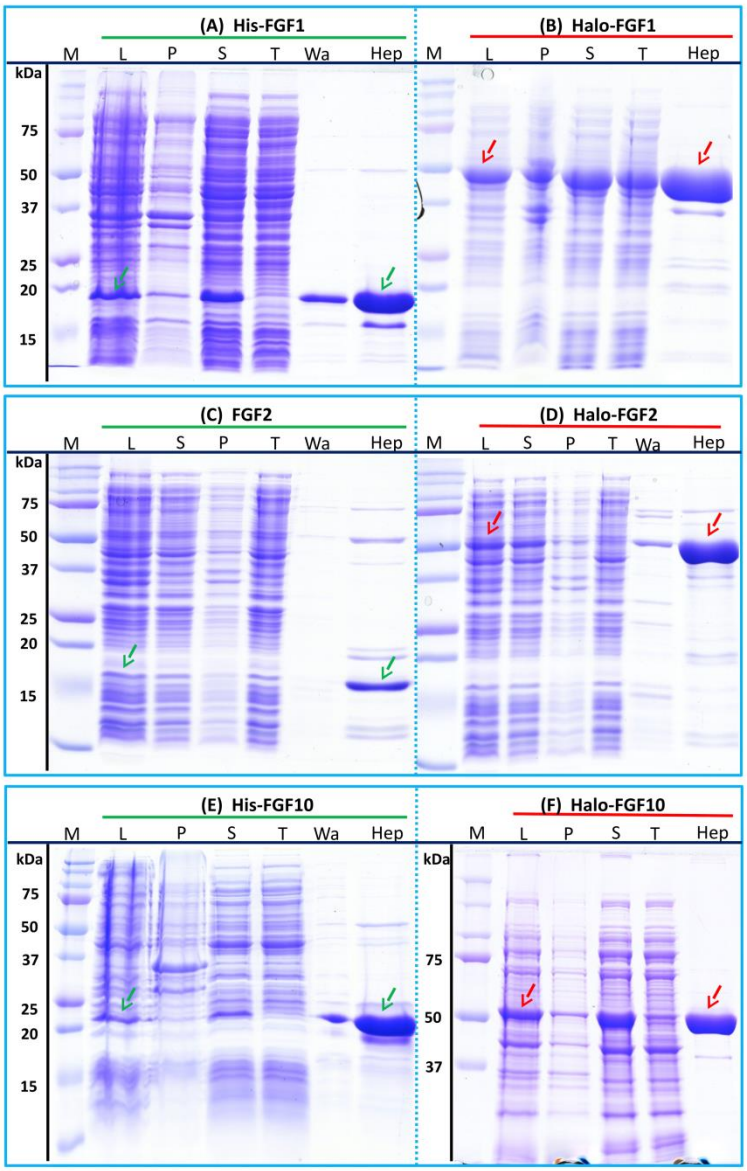


Figure 2 Expression and heparin affinity purification of His-FGF1, FGF2, His-FGF10, Halo-FGF1, Halo-FGF2 and Halo-FGF10. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-affinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Wa, wash of heparin-affinity column (Table 2); Hep, high NaCl eluate of heparin-affinity column (Table 2). Green arrows: FGF or His-FGF; red arrows: Halo-FGF.

Table 3 Summary of the molecular sizes and yields of His-FGFs and Halo-FGFs. The molecular weight of the proteins was calculated from their amino acid sequence. The concentrations and volumes of His-FGFs and Halo-FGFs recovered from heparin affinity chromatography were measured. The impurities identified by SDS-PAGE were quantified using ImageJ relative to the band corresponding to His-FGF and to Halo-FGF and the amount of protein in the eluate from heparin chromatography adjusted accordingly, to provide an estimate of the yield.

FGFs	Molecular weight (kDa)		Yield (mg/L)	
	HisTag	HaloTag	HisTag	HaloTag
FGF1	19.1	50.9	14	16
FGF2	17.3 No Tag	52.2	2.5 No Tag	11
FGF3	28.2	60.0	0.5	11
FGF6	22.3	54.1	n.d.[a]	27
FGF7	22.2	54.0	0.6	5.6
FGF8	25.7	57.5	n.d.[a]	1.7
FGF10	22.7	54.5	7.7	9.3
FGF16	26.9	58.7	n.d.[a]	1.0
FGF17	25.8	57.6	n.d.[a]	1.5
FGF20	26.9	58.6	n.d.[a]	10
FGF22	20.5	52.3	n.d.[a]	2.0

Notes.

[a] Not detected. Insufficient soluble protein for reliable quantification.

Transformation of SoluBL21with the plasmid encoding Halo-FGF7 yielded the expected number of colonies, indicating that the fusion protein was not toxic. Bands corresponding to the molecular size of Halo-FGF3 and Halo-FGF7 were observed in the cell lysates (Figs. 3B and 3D, lane L, red arrow) and in the soluble fraction obtained after centrifugation, whereas the pellet has relatively weaker bands (Figs. 3B and 3D, lanes P and S), indicating that Halo-FGF3 and Halo-FGF7 were soluble. Heparin chromatography of the soluble factions demonstrated that large amounts of Halo-FGF3 and Halo-FGF7 retained their heparin binding interaction with the polysaccharide (Figs. 3B and 3D, lane Hep).

The yields of Halo-FGF3 and of Halo-FGF7 were 21-fold and 9-fold greater than of the corresponding His-tagged FGF (Table 3). Thus, the presence of the HaloTag N-terminal fusion increased the amounts of FGF3 and FGF7 substantially, even after taking into account the larger size of these fusion proteins (Table 3).

Group 3: insoluble proteins

His-FGF6, His-FGF8, His-FGF22, His-FGF17, His-FGF16 and His-FGF20 were all expressed, albeit at different levels. After centrifugation, bands corresponding to the molecular sizes of these proteins were detected in the pellet (Fig. 4, compare lanes P and S, green arrow). Although small amounts of protein, such as bands corresponding to His-FGF6, His-FGF16 and His-FGF20, were observed in the supernatants (Fig. 4, lanes S), no protein were detected in the eluate from heparin chromatography, which might suggest these proteins were either small soluble aggregates or not properly folded. It has

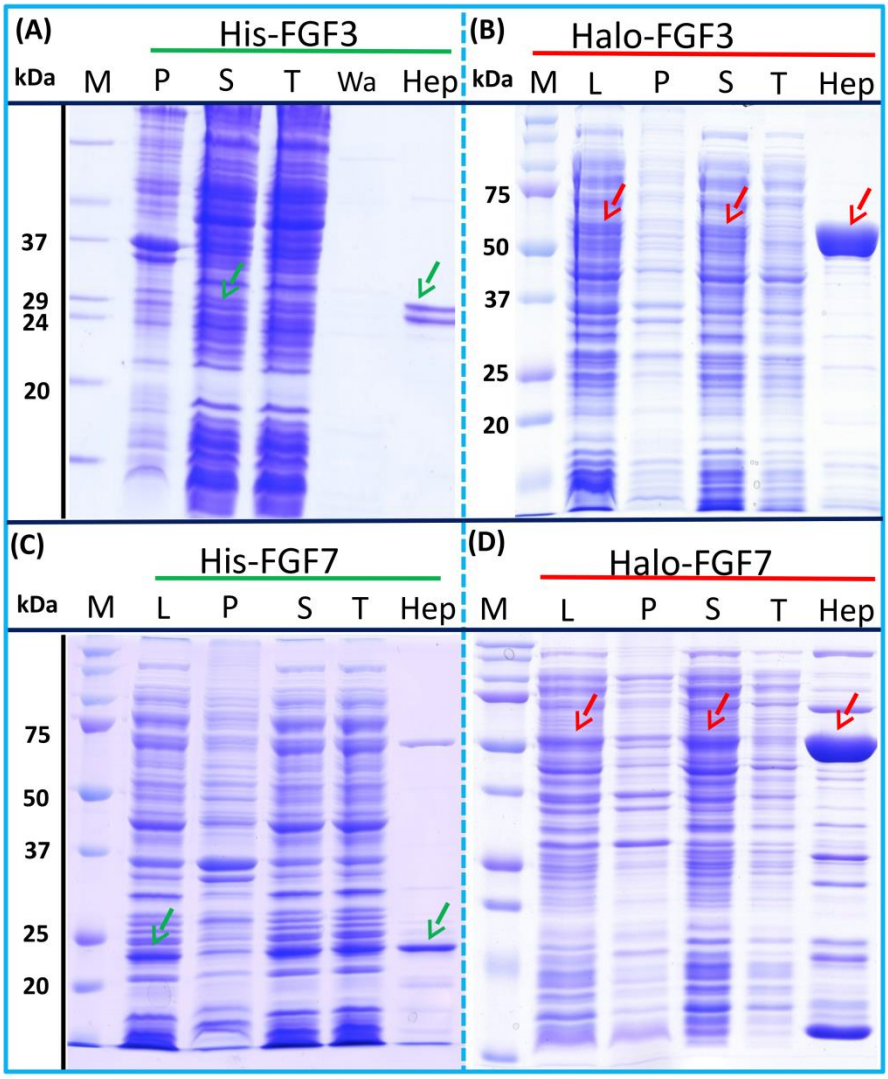


Figure 3 Expression and heparin binding-affinity chromatography of His-FGF3, His-FGF7, Halo-FGF3 and Halo-FGF7. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-afinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Hep, high [NaCl] eluate of heparin-affinity column (Table 2). Green arrows: His-FGF; red arrows: Halo-FGF.

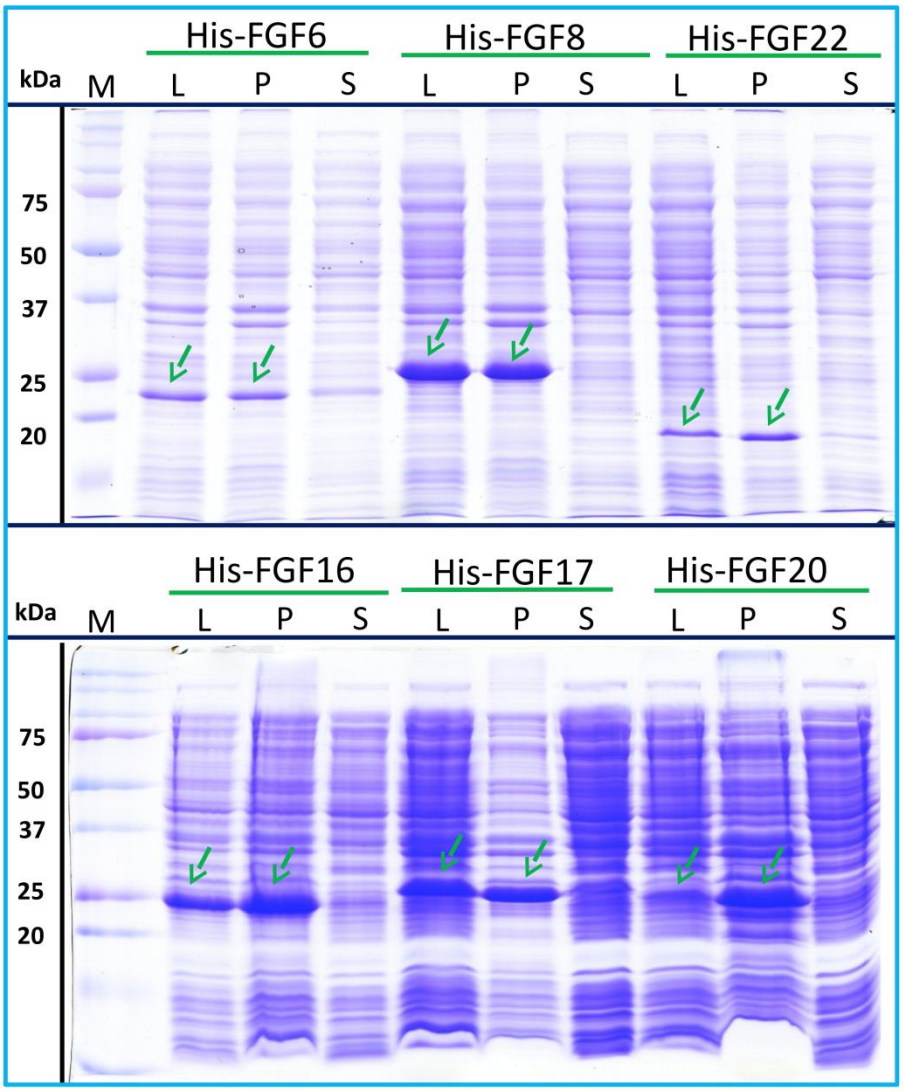


Figure 4 Expression test of His-FGF6, His-FGF8, His-FGF22, His-FGF17, His-FGF16 and His-FGF20. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The whole cell lysate, supernatant and pellet were analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant. Green arrows: His-FGF.

been reported that FGF20 could also be solubilised by high concentrations of arginine (*Maity, Karkaria & Davagnino, 2009*), which suggests that FGF20 in the lysis buffer has a tendency to aggregate. However, arginine would compete for binding of FGFs to heparin, which reduces the utility of this approach to solubilisation.

As illustrated by SDS-PAGE, all of the bands corresponding to the molecular size of Halo-FGF6, Halo-FGF8, Halo-FGF22, Halo-FGF17, Halo-FGF16 and Halo-FGF20 were clearly observed in the whole lysates, which suggested that all six proteins expressed well in *E. coli* (Fig. 5, lanes L, red arrow), particularly Halo-FGF6, Halo-FGF17, Halo-FGF16 and Halo-FGF20. Although some material corresponding to the expected molecular size of these Halo-FGFs was observed in the pellet after centrifugation of the cell lysates (Fig. 5, lanes P), there were strong bands corresponding to Halo-FGF6, Halo-FGF16 and Halo-FGF20 and weak bands corresponding to Halo-FGF8, Halo-FGF17 and Halo-FGF22 present in the soluble fractions (Fig. 5, lanes S). Following application to a heparin affinity column, most of Halo-FGF6 in the supernatant bound to heparin and was eluted by 1 M NaCl in Tris-Cl (Fig. 5A, lanes S, T and Hep). Halo-FGF8 Halo-FGF17 and Halo-FGF22 also bound to the heparin-affinity column reasonably efficiently, whereas a considerable amount of Halo-FGF16 and Halo-FGF20 did not bind (Figs. 5B–5E, lanes S and T). All four proteins could be recovered from heparin chromatography with high concentration NaCl-containing elution buffers (Table 2) (Figs. 5B–5E, lane Hep). When the Halo-FGF20 in the flow-through fraction (Fig. 5F, lane T) was applied to a second heparin-affinity chromatography column, a large amount of Halo-FGF20 was found to bind and could be eluted (Fig. 5F, lane Hep2). A considerable amount of Halo-FGF16 also failed to bind to the heparin affinity column (Fig. 5E, lane T), though the bound protein was eluted with high NaCl (Fig. 5E, lane Hep). This suggests that the capacity of the heparin affinity column for Halo-FGF20 was exceeded. The same explanation may underlie the presence of Halo-FGF16 in the flow-through fraction, though this protein was present at a slightly lower level. However, nothing is known about the preference of either FGF16 or FGF20 for binding structures in the polysaccharide, if so these were relatively rare in heparin, the column capacity might easily be exceeded. Alternatively, the Halo-FGF16 in the flow through fraction may represent protein that is in small aggregates and/or not properly folded.

Given that the amounts of soluble His-tagged FGF6, FGF8, FGF22, FGF17, FGF16 and FGF20 were not readily detectable, it is clear that the N-terminal HaloTag fusion significantly improved the expression of soluble protein. The yield of Halo-FGF6 and Halo-FGF20 was substantial (27 mg/L and 10 mg/L, respectively, Table 3). Although a lower yield of Halo-FGF8, Halo-FGF16, Halo-FGF17 and Halo-FGF22 (1 mg/L to 2 mg/L, Table 3) was obtained, it is nevertheless sufficient for many applications, including microscopy. However, the heparin affinity purification step did not produce entirely pure protein, as judged by coomassie staining (Figs. 2, 3 and 5).

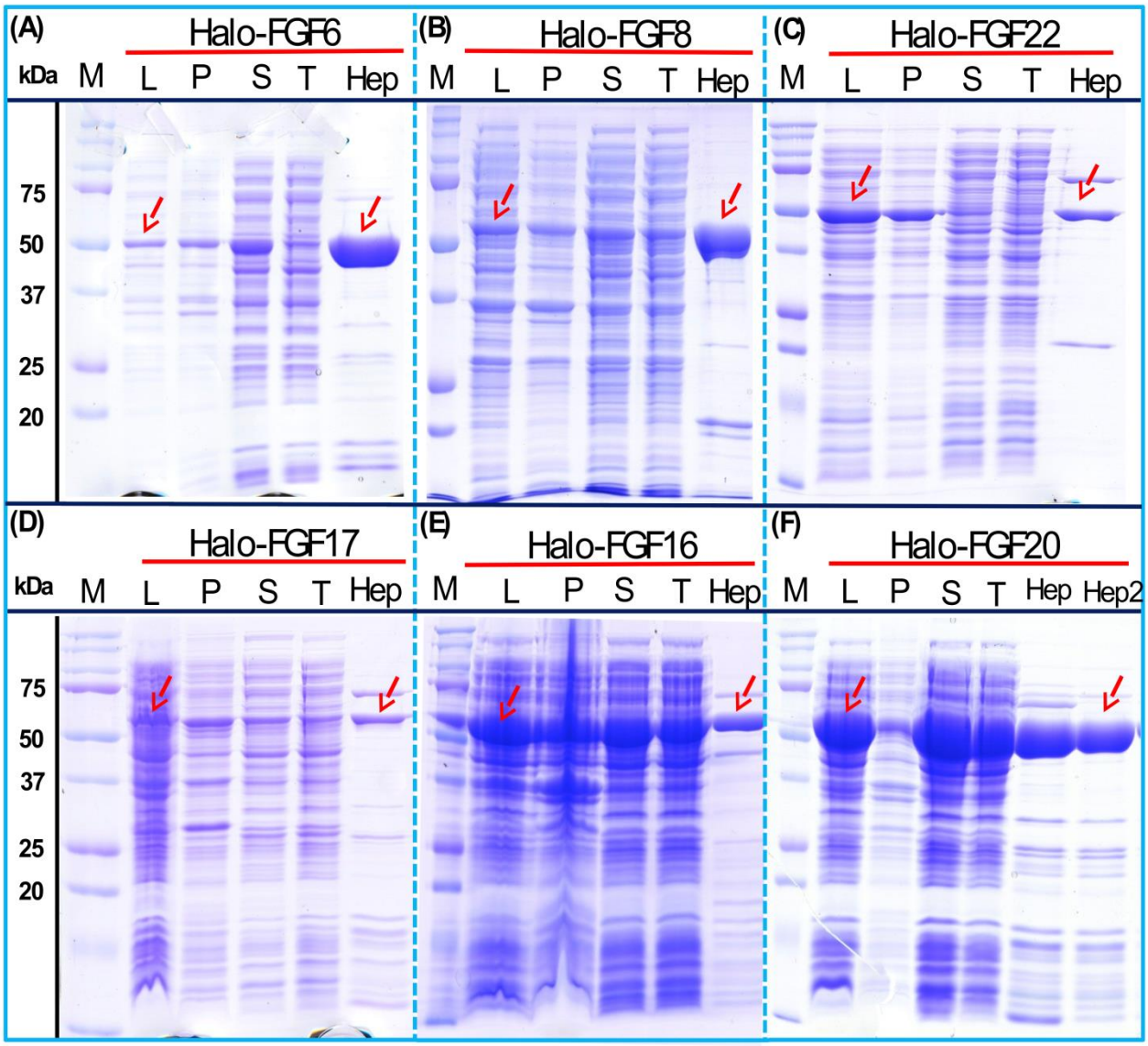


Figure 5 Expression and heparin affinity purification of Halo-FGF6, Halo-FGF8, Halo-FGF22, Halo-FGF17, Halo-FGF16 and Halo-FGF20. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-afinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Hep, high NaCl eluate of heparin-affinity column (Table 2). Red arrows: Halo-FGF.

Further purification of some Halo-FGFs

Four Halo-FGFs, Halo-FGF1, Halo-FGF7, Halo-FGF6 and Halo-FGF20 were chosen to determine whether the Halo-FGFs could be easily subjected to further purification, since there was clear evidence for impurities following heparin-affinity chromatography. Halo-FGF1 and Halo-FGF7 were successfully purified by Q anion-exchange chromatography (Figs. 6A and 6B, lane Q), which depends on the acidic isoelectric point of the HaloTag (pI: 4.77). For Halo-FGF6 and Halo-FGF20, advantage was taken of the acidic HaloTag and positive surfaces of FGFs, to enable a two-step ion-exchange purification of the eluate from heparin-affinity chromatography, using both DEAE anion and CM cation ion-exchange chromatography (Figs. 6C and 6D, lane DEAE and CM). The isolated Halo-FGFs are relatively pure, as is shown on the gels (Fig. 6).

Purification of FGFs by removing the HaloTag with TEV protease

The inclusion of a TEV site between the sequence of the HaloTag and FGF proteins provides a means to remove the HaloTag fusion partner in those instances where the HaloTag is not required for analysis (or when it may interfere with such analyses). Halo-FGF2 was first incubated with TEV protease to test whether the fusion protein could be cleaved by TEV. SDS-PAGE of the TEV digestion product of Halo-FGF2 shows that almost all of the protein was cleaved into the 35 kDa HaloTag (Fig. 7A, red arrow) and the 18 kDa FGF2 (Fig. 7A, green arrow). Thus, the cleavage site is fully accessible to TEV protease. Both Halo-FGF17 and Halo-FGF20 were also well digested by TEV protease and subsequently soluble FGF17 (Fig. 7B, green arrow) and FGF20 (Fig. 7C, green arrow) were purified by cation-exchange and heparin chromatography, respectively.

Most of FGF6 (Fig. 7D, lane W_{Dig}, green arrow) and FGF22 (Fig. 7F, lane W_{Dig}, green arrow) and a small proportion of FGF8 were also released from HaloTag (Figs. 7D–7F, lane W_{Dig} and S, red arrow), but these proteins were observed to aggregate upon cleavage. This suggested that these proteins were not very stable, at least in the buffer conditions used here, and required the HaloTag N-terminal fusion to remain soluble. The soluble FGF6 released by cleavage (Fig. 7D, lane S, green arrow) was applied to a heparin affinity column, but was observed to be concentrated at the top of the column where it formed a white aggregate. Very little protein was eluted with 1 M NaCl in PB buffer (Fig. 7D, lane E, green arrow). The disappearance of FGF8 and FGF22 in the soluble fractions after TEV digestion (Figs. 7E and 7F, lane S) showed that these two proteins were also not very soluble in the present buffer conditions without the HaloTag fusion partner.

Biological activities of FGFs and Halo-FGFs on Rama 27 fibroblasts and HaCaT keratinocytes

FGF1, FGF2 and FGF6 have a preference for FGFR1c (*Zhang et al., 2006a*), the predominant receptor expressed by Rama 27 fibroblasts (*Delehedde et al., 2000*; *Zhu et al., 2010*). When Rama 27 cells were stimulated with 25 pM FGF2 for 15 min, strong bands corresponding to dually phosphorylated $p44/42^{MAPK}$ were apparent (Fig. 8A), as observed previously (*Delehedde et al., 2000*; *Zhu et al., 2010*). Halo-FGF2 caused a similar

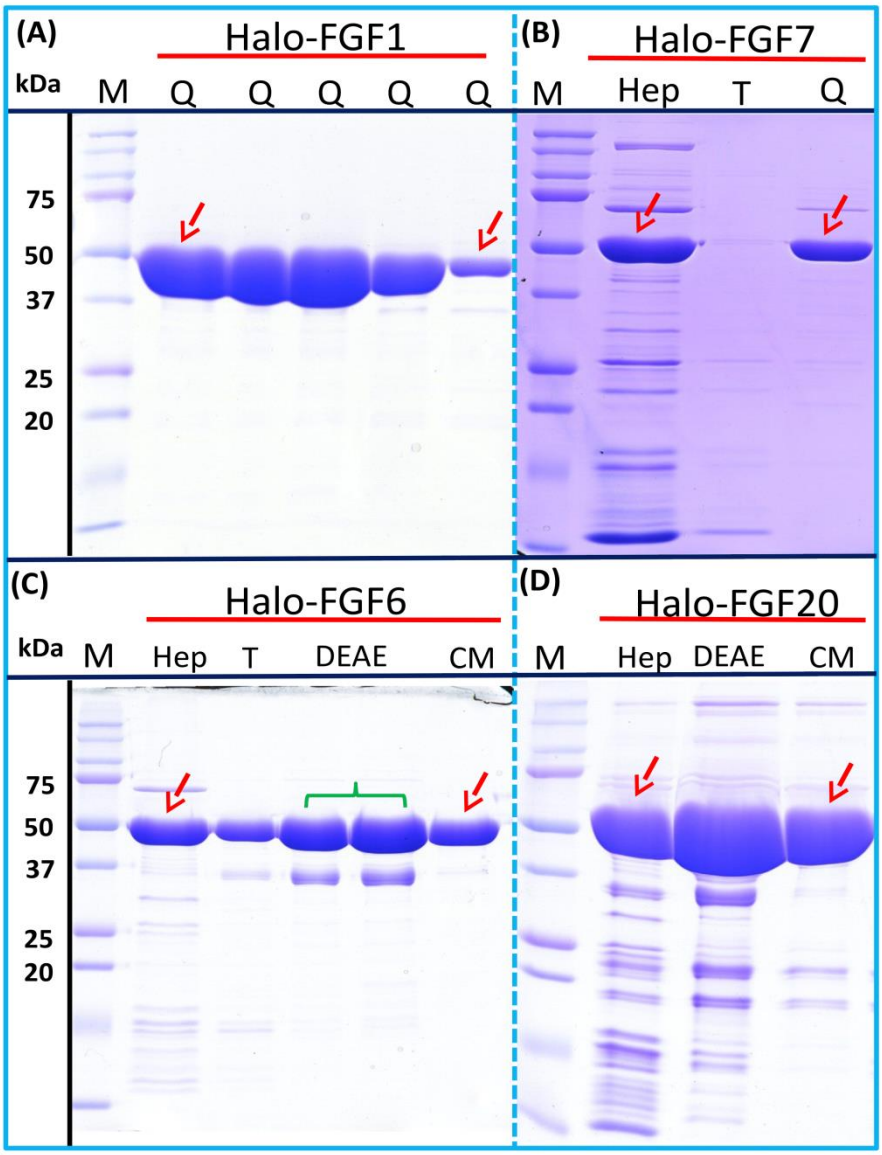


Figure 6 Further purification of the heparin affinity eluate of Halo-FGF1, Halo-FGF6, Halo-FGF7 and Halo-FGF20 by ion-exchange chromatography. The soluble Halo-FGF1 and Halo-FGF7 eluted from heparin chromatography was purified using Q ion-exchange chromatography, while CM and DEAE ion-exchange chromatography were used to purify Halo-FGF6 and Halo-FGF20. Lane M, markers; Hep, eluate from heparin chromatography, Figs. 2A, 3D, 5A and 5F; T, unbound, flow-through fraction from ion-exchange chromatography; Q, peak fractions collected from Q HP chromatography; DEAE eluate from DEAE chromatography, two identical samples; CM, eluate from CM chromatography. Red arrows: Halo-FGF.

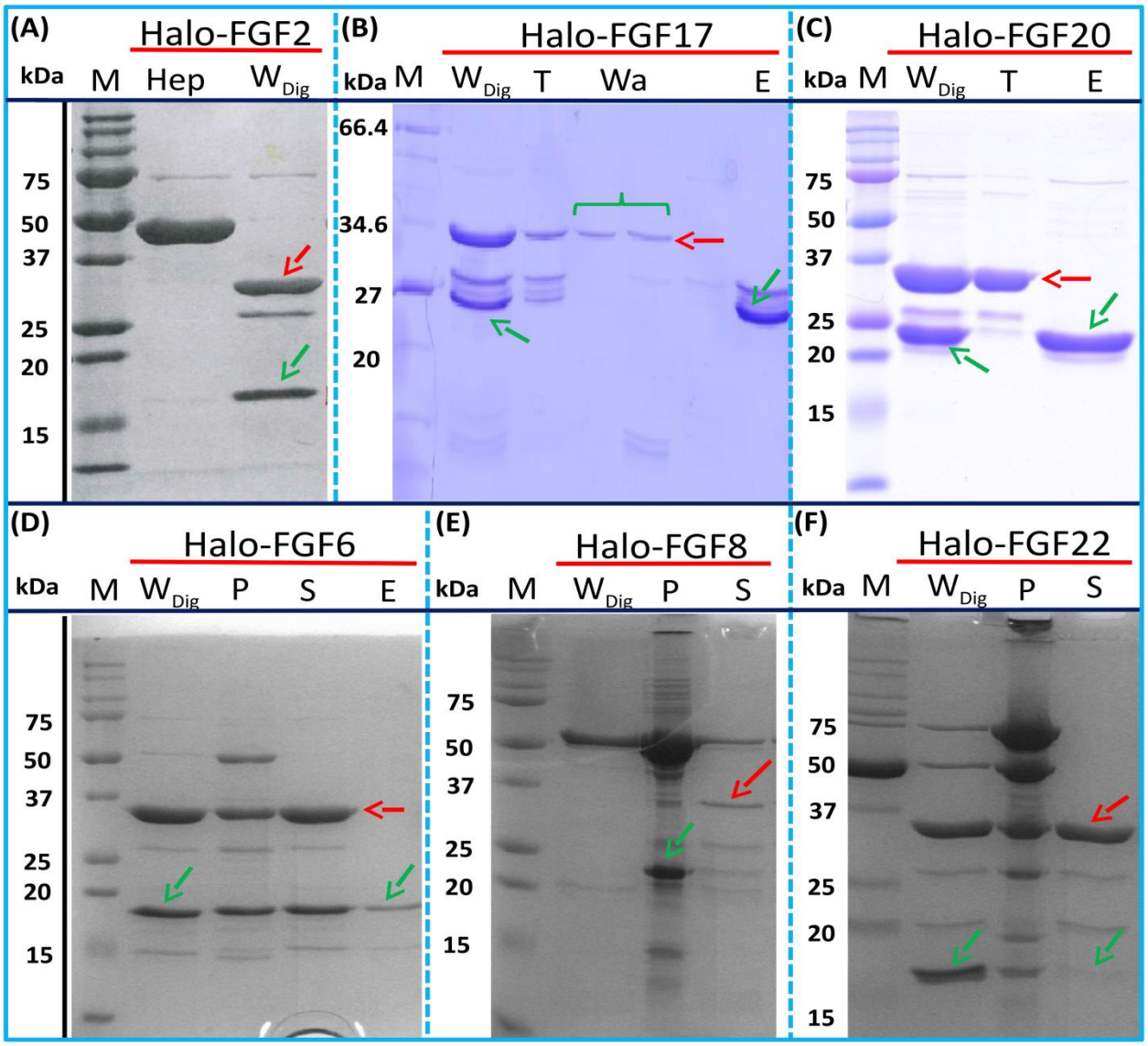


Figure 7 Cleavage of Halo-FGFs by TEV and purification. The eluates of Halo-FGF2, Halo-FGF17, Halo-FGF6, Halo-FGF8 and Halo-FGF22 from heparin-affinity chromatography and the Halo-FGF20 purified by heparin and ion-exchange chromatography were digested by TEV protease to separate the HaloTag and the FGF. Halo-FGF6, Halo-FGF8 and Halo-FGF22 became turbid after digestion and these samples were clarified by centrifugation. Then, the samples containing FGF6 and FGF20 were subjected to heparin chromatography and that of FGF17 to SP HP cation-exchange chromatography. Lanes M, markers; Hep, eluate from heparin chromatography; W_{Dig}, whole digestion product of Halo-FGFs purified by heparin chromatography; T, unbound, flow-through fraction from heparin chromatography; Wa, wash of SP HP cation-exchange chromatography; P, pellet following centrifugation of product of TEV digestion; S, supernatant after the centrifugation; E, high NaCl eluate of heparin or SP cation-exchange chromatography. Green arrows: FGF; red arrows: HaloTag.

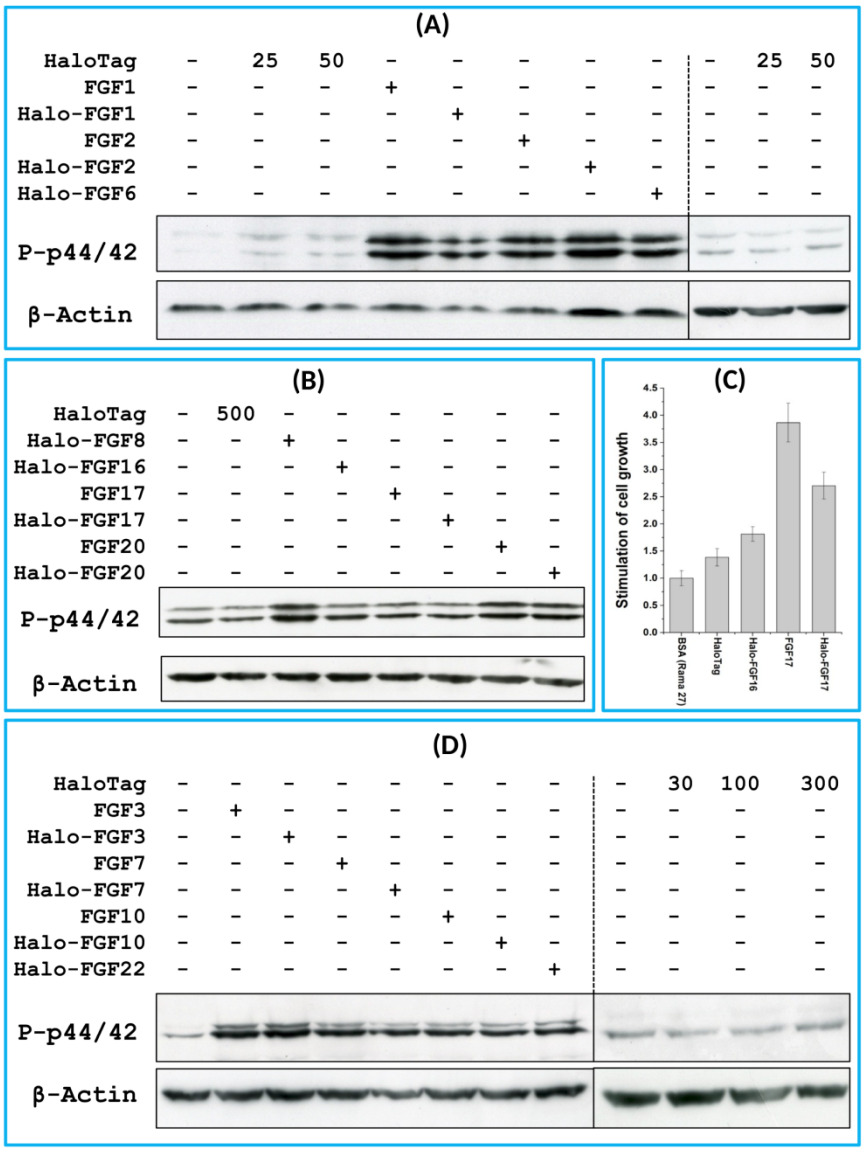


Figure 8 Activities of FGFs on Rama 27 fibroblasts and HaCaT keratinocytes. Cells were grown in 3 cm diameter dishes (Western blots) or 24-well plates growth assay, as described in the Materials and Methods. After incubation in SDM for 24 h (Rama 27) or 48 h (HaCaT), cells were stimulated with the FGF protein for 15 min (Western blot) or 68 h (cell growth assay). (A) Stimulation of p44/42^{MAPK} phosphorylation by 25 pM HaloTag, FGF2, Halo-FGF2 and 50 pM HaloTag, FGF1, Halo-FGF1 and Halo-FGF6 in Rama 27 fibroblasts. (B) Stimulation of p44/42^{MAPK} phosphorylation by 500 pM HaloTag, Halo-FGF8, Halo-FGF16, FGF17, Halo-FGF17, FGF20 and Halo-FGF20 in Rama 27 fibroblasts. (C) Stimulation of cell growth of Rama 27 fibroblasts by 10 nM Halo-FGF16, FGF17 and Halo-FGF17. (D) Stimulation of p44/42^{MAPK} phosphorylation by 300 pM His-FGF3, Halo-FGF3 and Halo-FGF22, 30 pM His-FGF7, Halo-FGF7, His-FGF10 and Halo-FGF10 in HaCaT cells.

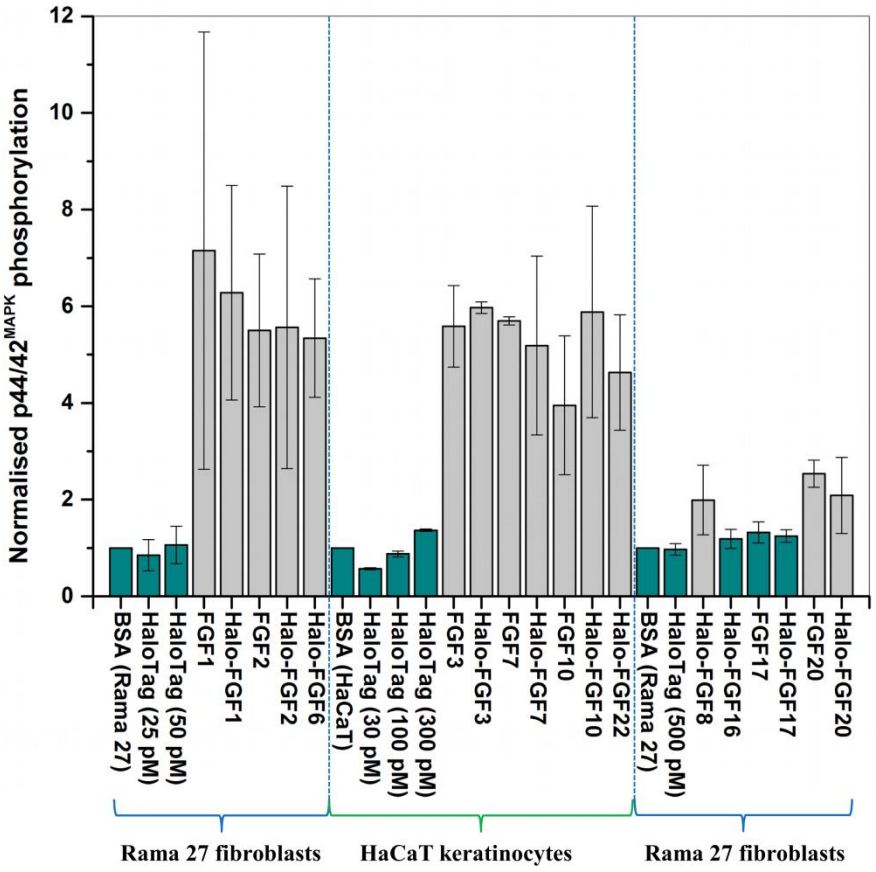


Figure 9 Quantification of p44/42^{MAPK} phosphorylation. The band intensities from two experiments were quantified with imageJ and normalised to the BSA control to compare the similarities and differences of stimulation of phosphorylation p44/42^{MAPK} by different FGFs. Results are the mean with the actual values from two independent experiments.

stimulation of phosphorylation of p44/42^{MAPK} (Fig. 8A). In contrast, the 25 pM or 50 pM HaloTag protein alone did not appreciably stimulate p44/42^{MAPK} phosphorylation. Therefore, the activity of Halo-FGF2 in this assay is equivalent to that of FGF2 (Fig. 9). In the case of FGF1, the N-terminal HaloTag also did not affect the ability of the growth factor to stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8A). FGF6 is not soluble without the HaloTag, so only the activity of the fusion protein could be tested, and it was found to stimulate the phosphorylation of p44/42^{MAPK} to an extent similar to that observed with FGF1 and FGF2 (Fig. 8A). Since FGF6 has the same receptor preference as FGF1 and FGF2 (*Zhang et al., 2006a*), this suggests Halo-FGF6 was fully active.

FGF8, FGF16, FGF17 and FGF20 have a preference for FGFR3c, but they are also able to activate FGFR1c, though higher concentrations of growth factor are required to elicit

activity (*Zhang et al., 2006a*). When 500 pM HaloTag was added to the cells, there was no detectable increase in phosphorylation of p44/42^{MAPK}, whereas Halo-FGF8, Halo-FGF20 and FGF20 at concentrations comparable to those used in previous work (*Zhang et al., 2006a*) were all found to stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8B). In contrast, Halo-FGF16, Halo-FGF17 and FGF17 did not cause a detectable increase in phosphorylation of p44/42^{MAPK} (Fig. 9). These data indicate that Halo-FGF8, FGF20 and Halo-FGF20 have biological activities on Rama 27 fibroblasts. The absence of stimulation of phosphorylation of p44/42^{MAPK} by Halo-FGF16 may reflect the fact that the ability of this FGF to activate FGFR1c is considerably lower than that of FGF8, FGF17 and FGF20 (*Zhang et al., 2006a*). However, the absence of stimulation of phosphorylation of p44/42^{MAPK} by FGF17 and Halo-FGF17 is more puzzling. One explanation may be that FGF16, and perhaps FGF17, do not cause the FGFR to activate strongly early biochemical signals that converge on p44/42^{MAPK}. To test this, the capacity of Halo-FGF16, Halo-FGF17 and FGF17 to stimulate cell growth was measured in Rama 27 fibroblasts. The results show that 10 nM HaloTag only weakly stimulated the growth of Rama27 fibroblasts. Halo-FGF16 caused the number of cells to double compared to the negative control, and this level was significantly ($p = 0.015$, Tukey test, OriginPro 9) above that observed in the presence of HaloTag alone (Fig. 8C). Halo-FGF17 and FGF17 were even more effective, as they caused a 3- to 4-fold increase in the number of cells (Fig. 8C). These results demonstrated that Halo-FGF16, FGF17 and Halo-FGF17 possess biological activities of similar potency as observed by others in growth assays (*Zhang et al., 2006a*).

The activity of members of the FGF7 subfamily were tested on HaCaT keratinocytes, as this cell type expresses the cognate receptor for these FGFs, FGFR2b (*Ron et al., 1993*). HaCaT cells have previously been shown to express more p42^{MAPk} than p44^{MAPk} (*Delehedde et al., 2002*). The data show clearly that HaloTag alone did not stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8D). In contrast, FGF3, FGF7 and FGF10, and the corresponding HaloTag fusion proteins stimulated p44/42^{MAPK} phosphorylation (Fig. 8C). FGF22, which is only soluble as a HaloTag fusion protein, also stimulated p44/42^{MAPK} phosphorylation to an extent similar to that seen with the other members of the subfamily (Fig. 8D). Thus, these Halo-FGFs retain full biological activity in this assay.

CONCLUSION

In this study, we identified four useful properties of N-terminal HaloTag fusions for the production of biologically active FGFs: (i) using the HaloTag can increase the yield of low expression FGFs, (ii) the HaloTag rendered FGF7 non-toxic; (iii) for the insoluble FGFs, the HaloTag enabled *E.coli* to express more soluble protein at low induction temperatures and maintain solubility during isolation and storage; (iv) a consequence of the low isoelectric point of HaloTag was that anion-exchange chromatography could be used as an orthogonal step in the purification of the Halo-FGFs. However, there are clearly limitations; for example, some of the FGFs did not retain solubility following cleavage from the HaloTag. This may reflect the fact that no single solubilisation tag is a universal panacea for resolving the problems of protein expression (*Costa et al., 2014*). Nevertheless, because

the HaloTag can enhance expression of soluble protein and provide a means to label FGF protein with different fluorescent dyes and quantum dots, e.g., *Los et al. (2008)*; *Zhang et al. (2006b)* it is clearly a versatile and useful tool for these two purposes and, therefore, worthwhile exploring as a part of experimental strategy with these aims.

ACKNOWLEDGEMENT

Xianqing Mao would like to thank Monika Dieterle for help with cloning Halo-FGF3.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The Cancer and Polio Research Fund and North West Cancer Research provided financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Cancer and Polio Research Fund and North West Cancer Research.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Changye Sun and Yong Li conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Sarah E. Taylor and Xianqing Mao performed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.
- Mark C. Wilkinson performed the experiments, reviewed drafts of the paper.
- David G. Fernig conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

REFERENCES

Asada M, Shinomiya M, Suzuki M, Honda E, Sugimoto R, Ikekita M, Imamura T. 2009. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochimica et Biophysica ACTA/General Subjects* **1790**:40–48 DOI 10.1016/j.bbagen.2008.09.001.

Beenken A, Mohammadi M. 2009. The FGF family: biology, pathophysiology and therapy. *Nature Reviews Drug Discovery* **8**:235–253 DOI 10.1038/nrd2792.

Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE. 1988. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *Journal of Cell Biology* **106**:761–771 DOI 10.1083/jcb.106.3.761.

Costa S, Almeida A, Castro A, Domingues L. 2014. Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Frontiers in Microbiology* **5**:63 DOI 10.3389/fmicb.2014.00063.

Danilenko DM, Montestruque S, Philo JS, Li TS, Hill D, Speakman J, Bahru M, Zhang MS, Konishi O, Itoh N, Chirica M, Delaney J, Hernday N, Martin F, Hara S, Talvenheimo J, Narhi LO, Arakawa T. 1999. Recombinant rat fibroblast growth factor-16: structure and biological activity. *Archives of Biochemistry and Biophysics* **361**:34–46 DOI 10.1006/abbi.1998.0967.

Delehedde M, Lyon M, Vidyasagar R, McDonnell TJ, Fernig DG. 2002. Hepatocyte growth factor/scatter factor binds to small heparin-derived oligosaccharides and stimulates the proliferation of human HaCaT keratinocytes. *Journal of Biological Chemistry* **277**:12456–12462 DOI 10.1074/jbc.M111345200.

Delehedde M, Seve M, Sergeant N, Wartelle I, Lyon M, Rudland PS, Fernig DG. 2000. Fibroblast growth factor-2 stimulation of p42/44(MAPK) phosphorylation and I kappa B degradation is regulated by heparan sulfate/heparin in rat mammary fibroblasts. *Journal of Biological Chemistry* **275**:33905–33910 DOI 10.1074/jbc.M005949200.

Duchesne L, Gentili D, Comes-Franchini M, Fernig DG. 2008. Robust ligand shells for biological applications of gold nanoparticles. *Langmuir* **24**:13572–13580 DOI 10.1021/la802876u.

Duchesne L, Octeau V, Bearon RN, Beckett A, Prior IA, Lounis B, Fernig DG. 2012. Transport of fibroblast growth factor 2 in the pericellular matrix is controlled by the spatial distribution of its binding sites in heparan sulfate. *PLoS Biology* **10**:e1001976 DOI 10.1371/journal.pbio.1001361.

Ferreira T, Rasband W. 2012. ImageJ user guide—Analyze: Gels. *Available at http://rsbweb.nih.gov/ij/docs/guide/146-30.html#toc-Subsection-30.13* (accessed 09 December 2014).

Huang ZH, Hwang P, Watson DS, Cao LM, Szoka FC. 2009. Tris-nitrilotriacetic acids of subnanomolar affinity toward hexahistidine tagged molecules. *Bioconjugate Chemistry* **20**:1667–1672 DOI 10.1021/bc900309n.

Itoh N. 2007. The FGF families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease. *Biological & Pharmaceutical Bulletin* **30**:1819–1825 DOI 10.1248/bpb.30.1819.

Jeffers M, McDonald WF, Chillakuru RA, Yang MJ, Nakase H, Deegler LL, Sylander ED, Rittman B, Bendele A, Sartor RB, Lichenstein HS. 2002. A novel human fibroblast growth factor treats experimental intestinal inflammation. *Gastroenterology* **123**:1151–1162 DOI 10.1053/gast.2002.36041.

Kalinina J, Byron SA, Makarenkova HP, Olsen SK, Eliseenkova AV, Larochelle WJ, Dhanabal M, Blais S, Ornitz DM, Day LA, Neubert TA, Pollock PM, Mohammadi M. 2009. Homodimerization controls the fibroblast growth factor 9 subfamily's receptor binding and heparan sulfate-dependent diffusion in the extracellular matrix. *Molecular and Cellular Biology* **29**:4663–4678 DOI 10.1128/MCB.01780-08.

Ke YQ, Fernig DG, Smith JA, Wilkinson MC, Anandappa SY, Rudland PS, Barraclough R. 1990. High level production of human acidic fibroblast growth factor in *Escherichia coli* cells inhibition of DNA synthesis in Rat mammary fibroblasts at high concentrations of growth factor. *Biochemical and Biophysical Research Communications* **171**:963–971 DOI 10.1016/0006-291X(90)90778-L.

Knoerzer W, Binder HP, Schneider K, Gruss P, Mccarthy JEG, Risau W. 1989. Expression of synthetic genes encoding bovine and human basic fibroblast growth-factors (bFGFs) in *escherichia-coli*. *Gene* **75**:21–30 DOI 10.1016/0378-1119(89)90379-X.

Lata S, Reichel A, Brock R, Tampe R, Piehler J. 2005. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *Journal of the American Chemical Society* **127**:10205–10215 DOI 10.1021/ja050690c.

Lin XH. 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* **131**:6009–6021 DOI 10.1242/dev.01522.

Loo BM, Salmivirta M. 2002. Heparin/heparan sulfate domains in binding and signaling of fibroblast growth factor 8b. *Journal of Biological Chemistry* **277**:32616–32623 DOI 10.1074/jbc.M204961200.

Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohane RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV. 2008. HatoTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology* **3**:373–382 DOI 10.1021/cb800025k.

Macarthur CA, Lawshe A, Xu JS, Santosocampo S, Heikinheimo M, Chellaiah AT, Ornitz DM. 1995. Fgf-8 isoforms activate receptor splice forms that are expressed in mesenchymal regions of mouse development. *Development* **121**:3603–3613.

Maity H, Karkaria C, Davagnino J. 2009. Effects of pH and arginine on the solubility and stability of a therapeutic protein (fibroblast growth factor 20): relationship between solubility and stability. *Current Pharmaceutical Biotechnology* **10**:609–625 DOI 10.2174/138920109789069297.

Ohana RF, Encell LP, Zhao K, Simpson D, Slater MR, Urh M, Wood KV. 2009. HaloTag7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification. *Protein Expression and Purification* **68**:110–120 DOI 10.1016/j.pep.2009.05.010.

Ornitz DM. 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* **22**:108–112 DOI 10.1002/(SICI)1521-1878(200002)22:2<108::AID-BIES2>3.0.CO;2-M.

Pizette S, Batoz M, Prats H, Birnbaum D, Coulier F. 1991. Production and functional-characterization of human recombinant FGF-6 protein. *Cell Growth and Differentiation* **2**:561–566.

Ron D, Bottaro DP, Finch PW, Morris D, Rubin JS, Aaronson SA. 1993. Expression of biologically-active recombinant keratinocyte growth-factor—structure-function analysis of amino-terminal truncation mutants. *Journal of Biological Chemistry* **268**:2984–2988.

Roullier V, Clarke S, You C, Pinaud F, Gouzer G, Schaible D, Marchi-Artzner V, Piehler J, Dahan M. 2009. High-affinity labeling and tracking of individual histidine-tagged proteins in live cells using Ni^{2+} Tris-nitrilotriacetic acid quantum dot conjugates. *Nano Letters* **9**:1228–1234 DOI 10.1021/nl9001298.

Rudland PS, Twiston Davies AC, Tsao SW. 1984. Rat mammary preadipocytes in culture produce a trophic agent for mammary epithelia-prostaglandin E2. *Journal of Cellular Physiology* **120**:364–376 DOI 10.1002/jcp.1041200315.

Smith JA, Winslow DP, Rudland PS. 1984. Different growth factors stimulate cell division of rat mammary epithelial, myoepithelial and stromal cell lines in culture. *Journal of Cellular Physiology* **119**:120–126 DOI 10.1002/jcp.1041190310.

Susumu K, Medintz IL, Delehanty JB, Boeneman K, Mattoussi H. 2010. Modification of poly(ethylene glycol)-capped quantum dots with nickel nitrilotriacetic acid and self-assembly with histidine-tagged proteins. *Journal of Physical Chemistry C* **114**:13526–13531 DOI 10.1021/jp103872j.

Tinazli A, Tang JL, Valiokas R, Picuric S, Lata S, Piehler J, Liedberg B, Tampe R. 2005. High-affinity chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. *Chemistry-a European Journal* **11**:5249–5259 DOI 10.1002/chem.200500154.

Turner N, Grose R. 2010. Fibroblast growth factor signalling: from development to cancer. *Nature Reviews Cancer* **10**:116–129 DOI 10.1038/nrc2780.

Uchinomiya S, Nonaka H, Fujishima S, Tsukiji S, Ojida A, Hamachi I. 2009. Site-specific covalent labeling of His-tag fused proteins with a reactive Ni(II)-NTA probe. *Chemical Communications* **39**:5880–5882 DOI 10.1039/b912025d.

Vogel A, Rodriguez C, IzpisuaBelmonte JC. 1996. Involvement of FGF-8 in initiation, outgrowth and patterning of the vertebrate limb. *Development* **122**:1737–1750.

Xu RY, Ori A, Rudd TR, Uniewicz KA, Ahmed YA, Guimond SE, Skidmore MA, Siligardi G, Yates EA, Fernig DG. 2012. Diversification of the structural determinants of fibroblast growth factor-heparin interactions implications for binding specificity. *Journal of Biological Chemistry* **287**:40061–40073 DOI 10.1074/jbc.M112.398826.

Yu S, Burkhardt M, Nowak M, Ries J, Petrásek Z, Scholpp S, Schwille P, Brand M. 2009. FGF8 morphogen gradient is formed by a source–sink mechanism with freely diffusing molecules. *Nature* **461**:533–536 DOI 10.1038/nature08391.

Zhang XQ, Ibrahimi OA, Olsen SK, Umemori H, Mohammadi M, Ornitz DM. 2006a. Receptor specificity of the fibroblast growth factor family— the complete mammalian FGF family. *Journal of Biological Chemistry* **281**:15694–15700 DOI 10.1074/jbc.M601252200.

Zhang Y, So MK, Loening AM, Yao HQ, Gambhir SS, Rao JH. 2006b. HaloTag protein-mediated site-specific conjugation of bioluminescent proteins to quantum dots. *Angewandte Chemie-International Edition* **45**:4936–4940 DOI 10.1002/anie.200601197.

Zhu HY, Duchesne L, Rudland PS, Fernig DG. 2010. The heparan sulfate co-receptor and the concentration of fibroblast growth factor-2 independently elicit different signalling patterns from the fibroblast growth factor receptor. *Cell Communication and Signaling* **8**:14 DOI 10.1186/1478-811X-8-14.

3.3 Supplemental results of expressions and purification of FGFs

3.3.1 Bacteria transformation

Seventy ng of pET-14b-FGF-2, pET-14b-FGF-2 (mutants), pETM-11-FGF-3, pETM-11-FGF-4, pETM-11-FGF-16 were transformed into C41 (DE3) cell, while pETM-11-FGF-7 and pET-14b (halo)-FGF-17 were transformed into 70 µL BL21 (DE3) plysS cells (Section 2.3.3).

3.3.2 Bacterial cultures

3.3.2.1 His-tagged FGF3, His-tagged FGF7 and His-tagged FGF16

C41 cells containing FGF2 mutants and BL21 cells containing His-tagged FGF3, His-tagged FGF7 or His-tagged FGF16 were cultured into 3 L (6x 500 mL) LB broth with kanamycin (50 mg/mL) and induced overnight with IPTG (1 mM) at 16°C (180 rpm) (section 2.4.6).

3.3.2.2 FGF-2 and its mutants

C41 cells containing FGF-2 and its mutants were cultured in 3 L (6x 500 mL) LB broth- ampicillin (10 mg/mL) and induced with IPTG (1 mM) at 37°C (240 rpm) for 3 h or at16°C (Mutants) (180 rpm) overnight (section 2.4.6).

3.3.2.3 His-tagged FGF4

C41 cells containing His-tagged FGF4 were cultured in 3 L (6x 500 mL) 2x YT medium- kanamycin and induced with IPTG (1 mM) at 37°C (240 rpm) for 3 h (section 2.4.6).

3.3.3 Results of purification of FGFs

The purification of His-tagged FGF3, His-tagged FGF7, His-tagged FGF16 and Halo-tagged FGF-17 has been described in Section 3.2. The purification of FGF4, FGF2 and its mutants is described below. In all cases, cells were collected, lysed and the soluble proteins were isolated by centrifugation as described in [171].

3.3.3.1 FGF2

After cell breakage and centrifugation, the supernatant was loaded onto a 1 mL Hi-Trap heparin column and eluted with buffer (50 mM Tris-Cl, 1.0 M NaCl, pH 7.2). Some contaminants were apparent in the eluate (Fig. 3.2), however, so further purification was required. The eluate was diluted 6-fold with buffer (50 mM Tris-HCl, pH 7.2) and applied to a 1 mL HiTrap SP HP cation-exchange column, which was then developed with a gradient of 150 mM to 1 M NaCl, both in 50 mM Tris-Cl, pH 7.2). The fractions containing pure FGF2 (Fig. 3.2) were stored at -80°C.

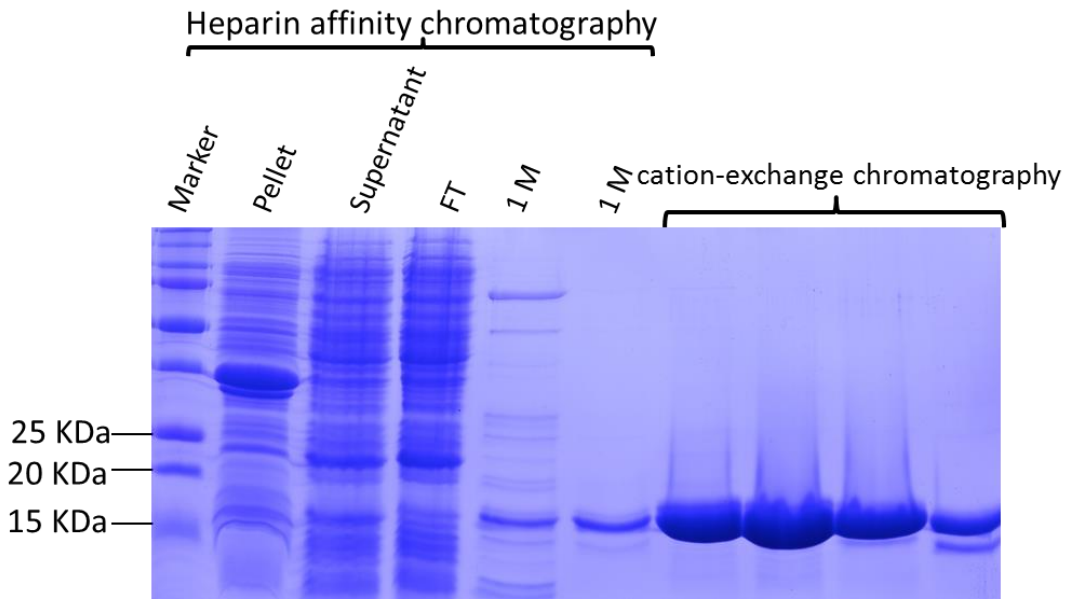


Figure 3.2 SDS PAGE of different stage of FGF2 purification (Mw≈ 16.7 kDa).

A 3 L culture was collected by centrifugation, disrupted by sonication and cell debris removed by further centrifugation. The soluble proteins were applied onto a 1 mL Hi-Trap Heparin column. The 1 M NaCl eluate was diluted 6-fold and then applied to 1 mL Hi-Trap SP HP cation-exchange column. The fraction containing FGF2 was eluted with 50 mM Tris-HCl in 1 M NaCl, pH 7.2. 'Pellet', - insoluble part of sample following centrifugation of disrupted bacteria; 'Supernatant', - soluble part of disrupted bacteria; 'FT', - unbound material during sample loading and buffer washing of the Hi-Trap heparin column; '1 M', - elution with 1 M NaCl in 50 mM Tris buffer pH 7.2; 'Ion-exchange chromatography', - the fractions eluted from Hi-Trap SP HP column.

3.3.3.2 FGF2 mutant (HBS2)

The expression and purification of FGF2 (HBS2) was performed similarly to that of FGF2. After cell breakage, the supernatants were loaded on a heparin affinity column. However, FGF2 (HBS2) did not elute in 1 M NaCl but instead was found to elute in 1.5 M NaCl, 50 mM Tris-Cl, pH 7.2 (Fig. 3.3A). The eluant from the heparin affinity chromatography was diluted to less than 150 mM NaCl with 50 mM Tris-Cl, pH 7.2 and then applied to and eluted from a cation-exchange column, exactly as for FGF2 (Fig. 3.3B). Two bands highlighted by black arrow and red arrow on gel were cut and digested with trypsin. The digested peptides were applied to MALDI-Q-TOF and identified as part of FGF2 (HBS2) (data not shown). These two bands could be FGF2 (HBS2) with or without histag.

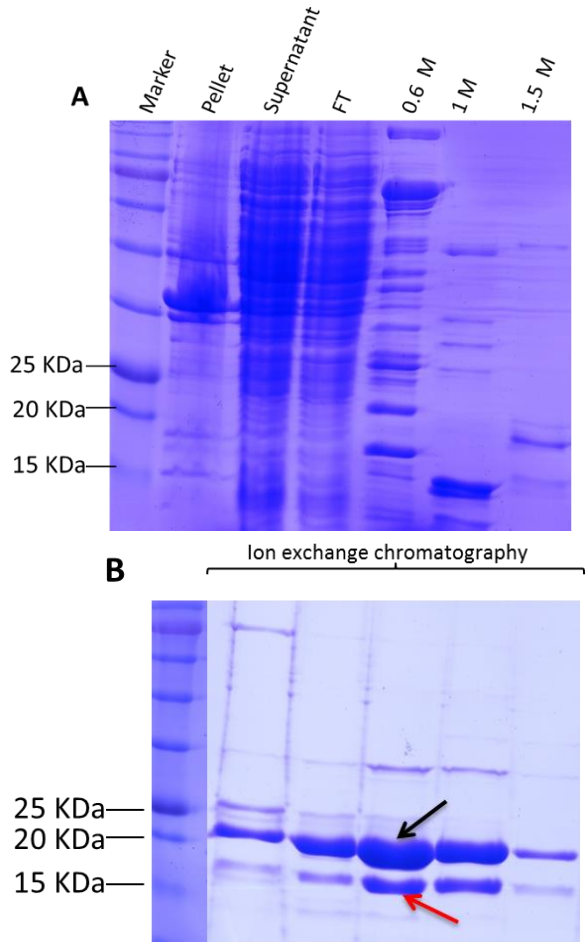


Figure 3.3 SDS PAGE of samples containing FGF2 (HBS2) (Mw≈ 18.7 kDa).

Sample of disrupted bacteria was applied to a 1 mL Hi-Trap heparin HP column. The 1 M NaCl eluate was diluted 6-fold with 50 mM Tris, pH 7.2 and then applied to a 1 mL Hi-Trap SP HP cation-exchange column. The fraction containing FGF2 (HBS2) was eluted with 50 mM Tris-Cl, 1 M NaCl, pH 7.2. *A*, heparin affinity chromatography, 'Pellet', - insoluble part of sample followed by centrifugation of disrupted bacteria; 'Supernatant', - soluble part of disrupted bacteria; 'FT', - flow through and wash fractions from the heparin-affinity column; '1 M', - fractions eluted with 1 M NaCl, 50 mM Tris, pH 7.2; *B*, 'cation-exchange chromatography', - the eluate from Hi-Trap SP HP column.

3.3.3.3 FGF2 mutant (HBS3)

Following cell breakage, the supernatant was applied to a 1 mL Hi-Trap heparin HP column and washed with buffer Bh (50 mM Tris-Cl, 0.6 M NaCl, pH 7.2). The fraction containing FGF2 (HBS3) was eluted with elution buffer Ch (50 mM Tris-Cl, 1.0 M NaCl, pH 7.2) (Fig 3.4).

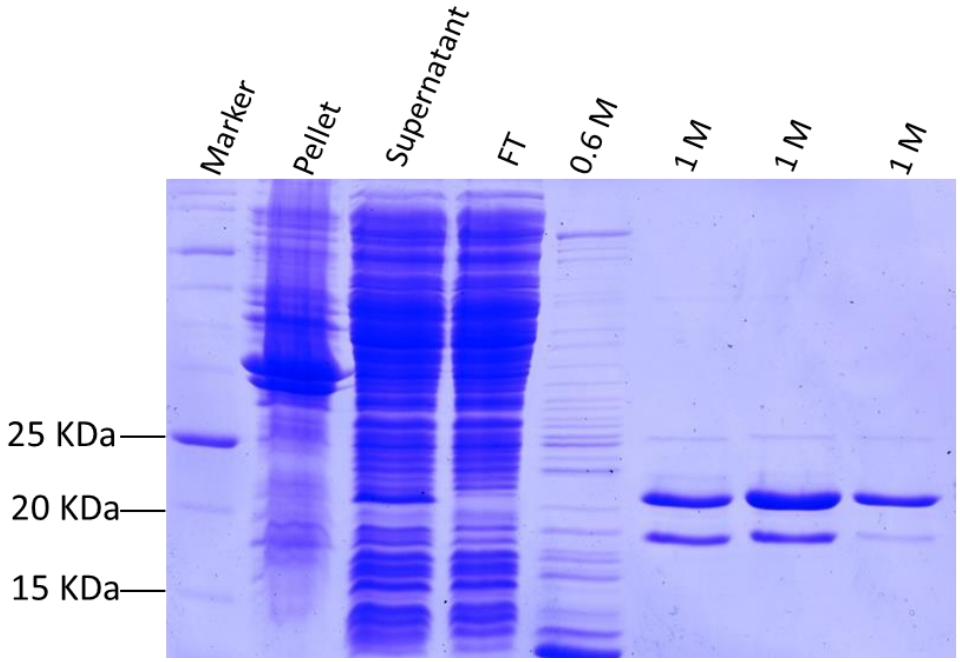


Figure 3.4 SDS PAGE of samples containing FGF2 (HBS3) (Mw≈ 21 kDa).

A 3 L culture was disrupted by sonication and the soluble part was subjected to heparin-affinity chromatography. The fraction contained FGF2 (HBS3) was eluted with 1 M NaCl, 50 mM Tris-HCl, pH 7.2. 'Pellet', - insoluble part of sample followed by centrifugation of disrupted bacteria; 'Supernatant', - soluble part of disrupted bacteria; 'FT', - fractions came from column during sample loading and buffer washing; '0.6 M', - fraction eluted with Tris buffer contained 0.6 M NaCl; '1 M', - fractions eluted with Tris buffer contained 1 M NaCl.

3.3.3.4 FGF4

Following cell breakage, the supernatant was applied to a 1 mL Hi-Trap heparin HP column, washed with buffer (0.6 M NaCl in 50 mM Tris-Cl, pH 7.2.), then eluted with buffer (1 M NaCl in 50 mM Tris-Cl, pH 7.2). To remove the contaminants in the fraction eluted from heparin (Fig. 3.5A), cation-exchange chromatography was used. The fraction was diluted 6-fold with buffer (50 mM Tris-Cl, pH 7.2) and applied to a 1 mL Hi-Trap SP HP column. The column was then mounted onto an AKTA system and eluted with a gradient of 0.15 M to 1 M NaCl in 50 Tris-Cl, pH 7.2 (Fig. 3.5B). The fractions containing pure FGF4 were stored at -80°C.

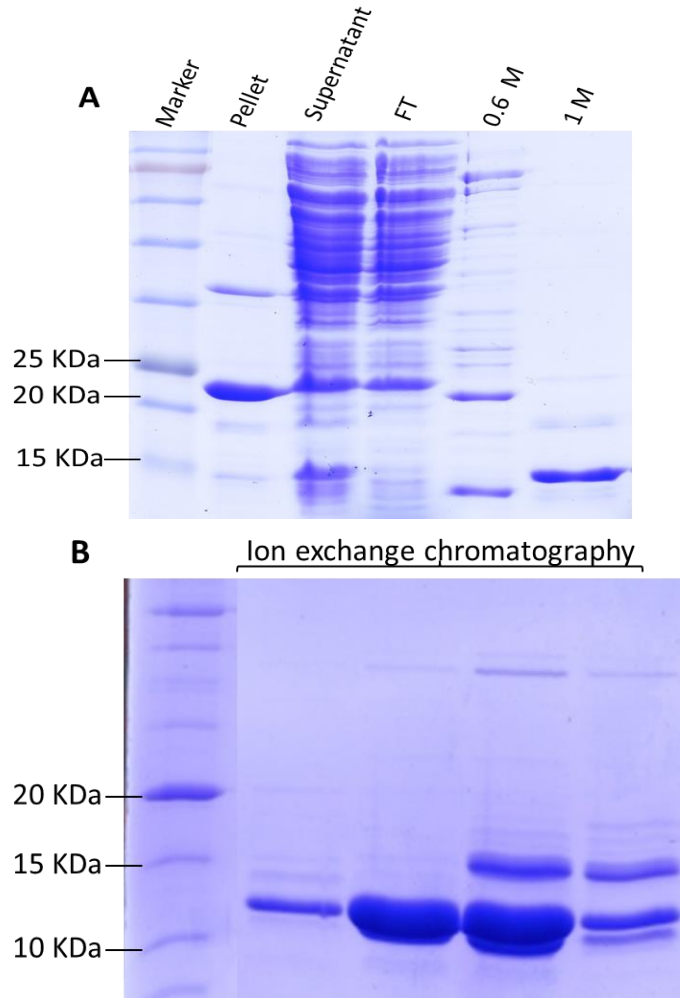


Figure 3.5 SDS PAGE of different stage of FGF4 purification (Mw≈ 13.7 kDa).

A 3 L culture was disrupted by sonication and the soluble part of disrupted bacteria was applied to a 1 mL Hi-Trap Heparin column. The 1 M NaCl eluate was diluted 6-fold and then applied to a 1 mL Hi-Trap SP HP column. The fraction containing FGF4 was eluted with a gradient of NaCl. *A*, 'Pellet', - insoluble part of sample following centrifugation of disrupted bacteria; 'Supernatant', - soluble part of disrupted bacteria; 'FT', - flow through fractions during sample loading and buffer wash; '1 M', - fractions eluted with 1 M NaCl in 50 mM Tris, pH 7.2; *B*, 'cation-exchange chromatography', - the fractions eluted from the Hi-Trap SP HP column.

3.4 Size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS)

SEC-MALLS was used to determine an accurate measurement of the molar mass of FGFs and whether the FGFs existed as monomers or oligomers.

After purification, FGF2, Histag-FGF2 mutants (HBS2, HBS3), Histag-FGF4, Halotag-FGF6, Histag-FGF10, FGF-17, and Histag-FGF20 were loaded onto the SEC-MALLS system and analysed using the on-board ASTRA software (version 6.1). The stability and accuracy of the system were first measured with a 50 µg bovine serum albumin standard. Afterwards, at least 50 µg FGF2 was analysed and its molecular weight and differential reflective index were obtained (Fig. 3.6A). The single peak and molecular mass (17.6 kDa) indicated that FGF2 existed as monomer. However, with FGF20 the molecular mass (39.8 kDa) was double that expected of a FGF20 monomer, which suggested that FGF20 is dimer (Fig. 3.6B). Since only one peak appeared during the measurement, this indicated that FGF20 exists as dimer and that no monomer-dimer equilibrium was apparent under these conditions (Fig. 3.6B).

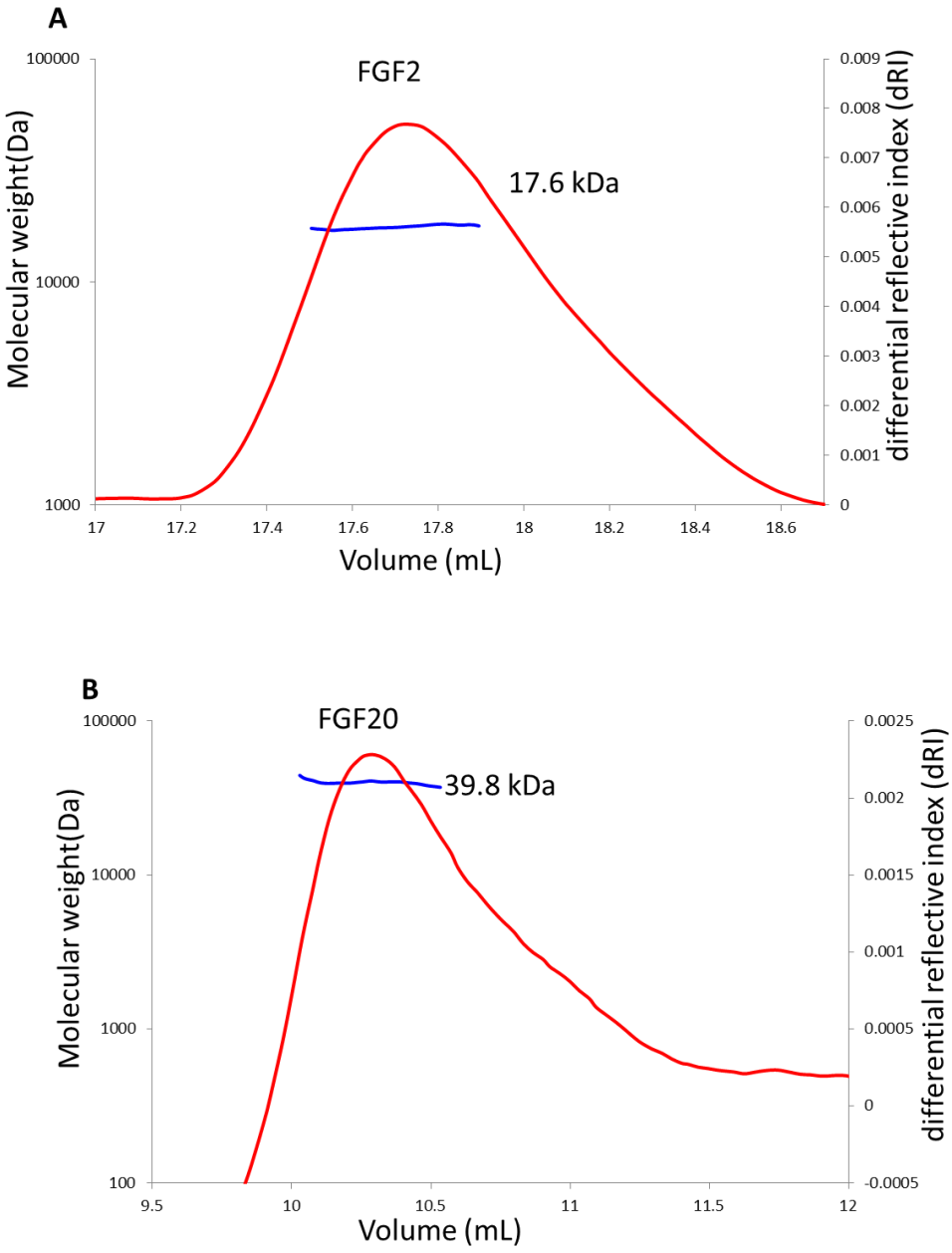


Figure 3.6 The average molecular mass per volume unit volume and the differential refractive index of FGF2 and FGF20. The red line shows the different reflective index (dRI). The blue line indicates the molecular mass of FGFs.

The FGF2 mutants where basic amino acids of heparin binding sites were mutated to alanine were measured and analysed by SEC-MALLS. The results show there was little difference in the molecular mass between the FGF2 mutants and the natural FGF2: FGF2 HBS2 (18.7 kDa) and FGF2 HBS3 (21 kDa), though there was a degree of peak broadening for FGF2 HBS3 (Figs 3.7A and 3.7B). This indicates that the mutation of the basic residues did not alter the oligomeric state of the FGF2. The FGF4 (13.7 kDa), FGF10 (kDa), Halotag-FGF6 (56.4 kDa) and FGF17 (23.2 kDa) were also analysed and the results show that these are all present as monomers (Fig. 3.7).

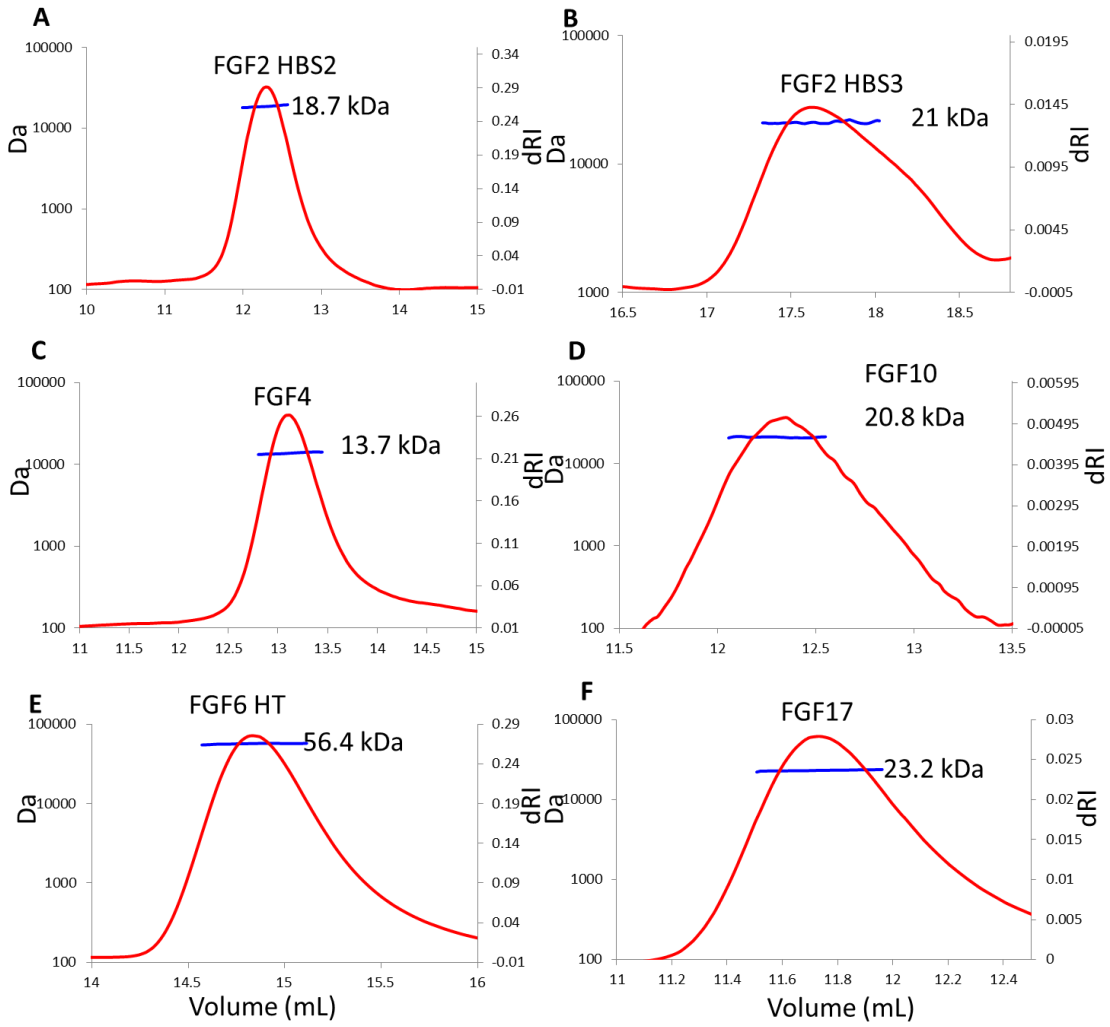


Figure 3.7 The average molecular mass per volume unit and the differential refractive index of FGFs (FGF2 mutants, FGF4, FGF10, Halotag-FGF6 and FGF17). The red line shows the different reflective index (dRI). The blue line indicates the molecular mass of FGFs.

Chapter 4 Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification

In Chapter 4, FGF-3, FGF-4, FGF-6, FGF-10, FGF-17, FGF-20, have been used as a defined system subjected to natural selection, to determine the specificities of their interactions with HS. This was measured from two perspectives. First, the heparin structures required for binding by these FGFs was measured by differential scanning fluorimetry, using a library of modified heparins and model glycosaminoglycans. Second, the extent of the primary canonical heparin binding site and the types of secondary heparin binding sites on these FGFs were determined by a lysine-targeted protect and label approach. By extending the analyses to these 6 paracrine FGFs, it will then be possible to combine the data with previous ones [34] and so provide a reasonable coverage of the FGF family: at least two FGFs from each subfamily. In this way it should be possible to determine the extent to which there is specificity and selectivity in the interactions of FGFs with HS and whether these correlate with the FGFR isoform binding properties of the FGFs and their phylogeny based on amino acid sequence, which is a major aim of this thesis (Section 1.9). This work will also provide insight into the broader question of specificity of the interactions of proteins with heparin and HS and provides the first set of data that will enable a later (Chapter 7) critical evaluation of what sulfated sequences of saccharides in HS actually mean from the perspective of protein binding.

4.1 Paper: Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification

Y. Li, C. Sun, E. A. Yates, M. C. Wilkinson, and D. G. Fernig, 'Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification, Open Biol, Submited (2015).

Contributions

Yong Li: Contributed to the design of the experiments, performed protein production, DSF, SEC-MALLs and MALDI-TOF, data analysis, drafted and wrote the manuscript.

Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification

Yong Li, Changye Sun, Edwin A. Yates, Mark C. Wilkinson, David G. Fernig

Department of Biochemistry, Institute of Integrative Biology, Biosciences Building, University of Liverpool, Crown Street, Liverpool L69 7ZB, United Kingdom

Corresponding Author:

David G. Fernig, Department of Biochemistry, Biosciences Building, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK. Tel.: +441517954471; Fax: +441517954406; E-mail: dgferning@liverpool.ac.uk.

Abstract

Heparan sulfate (HS) is key to the regulation of the functions of a large number (883) of extracellular regulatory proteins. However, the level of specificity of the protein-polysaccharide interactions and whether such specificity is of functional significance is still debated. As a family, the fibroblast growth factors (FGF) bind HS and they have a clear phylogenetic relationship at the level of amino acid sequence, which is mirrored in their functions. We have used members of the FGF family (FGF-3, FGF-4, FGF-6, FGF-10, FGF-17, FGF-20), as a defined system subjected to natural selection, to determine the specificities of their interactions with HS. Using a library of modified heparins and other glycosaminglycans, differential scanning fluorimetry and a protect and label strategy, it is clear that the FGFs from the same subfamily selectively bind oligosaccharide structures of similar sulfation pattern and length, whereas FGFs from different subfamilies have more divergent selectivity. Moreover, analysis of polysaccharide binding sites on the FGFs demonstrates that the presence and type of secondary binding sites is similar in FGFs from the same subfamily, but differs between subfamilies. The results demonstrate that specificity and selectivity underlies the interactions between FGFs and HS. Moreover, the data show that the molecular basis of the binding of FGFs to HS is likely to have been subjected to the same selection pressures that drove the expansion and divergence of the FGF family and their selectivity for particular isoforms of their receptor tyrosine kinase during the evolution of complex animals.

Keywords: Heparan sulfate (HS), fibroblast growth factor (FGF), specificity, heparin binding site;

Introduction

The glycosaminoglycan heparan sulfate (HS) regulates many aspects of cell communication by means of binding to over 883 extracellular proteins and thereby controlling their activities [113, 114, 156, 172, 173]. Two classic examples are the activation of antithrombin III by the polysaccharide, which contributes to the regulation of coagulation [174] and the control of the transport and effector functions of the paracrine fibroblast growth factors (FGFs) by their binding HS [119, 120]. A major challenge is to understand the structural basis of the interactions of proteins with HS and to what extent any molecular specificity and selectivity of these interactions is of functional significance.

HS consists of repeating disaccharide units joined by 1-4 linkages. The HS chains are always synthesized attached to a core protein to form heparan sulfate proteoglycans (HSPGs). It is the core protein that directs the HS chains to their functional location, which can be the cell surface or the extracellular matrix. Heparin, often used as an experimental proxy for HS on account of its underlying structural similarity, nevertheless is a more sulfated structure. The repeating units of HS consist of a glucuronic acid (GlcA) or its C5-epimer IdoA, and D-glucosamine (GlcN). The glucosamine may be N-acetylated (GlcNAc), N-sulfated (GlcNS), or unsubstituted (GlcN). The sugar chain is modified by the N-deacetylase / N-sulfotransferases (NDSTs) to replace the N-acetyl group of glucosamine with a sulfate group [115, 175, 176] and may be followed by C5 epimerization with C5 epimerase [107] and O-sulfation with 2-O, 3-O and the 6-O sulfotransferases [108, 109]. Since all the other modifications depend on the presence of N-sulfated glucosamine, the result is that HS chains have a domain structure: NA domains of no sulfation structure, NS domains of highly sulfated structures, and NA/NS domains comprising mixed disaccharides of GlcNAc and GlcNS [175, 177]. The sulfated structures are of functional significance, forming the protein binding domains [113].

The FGF family of 22 proteins has been divided into 7 subfamilies by phylogenetic analysis [29, 31]. Based on their action mechanism, FGFs can be classified into three types: intracrine, paracrine and endocrine. Only the paracrine FGFs bind to HS. Evidence for control by HS of FGF transport comes from a variety of experimental system. For example, mutations in gene encoding sugarless (sgl) and sulfateless (sfl), which are indispensable in the process of heparan sulfate chain biosynthesis, were identified as producing similar phenotypes to Wingless (Wg) or Hedgehog (Hh) signalling mutants [115]. Interactions with HS, which are occurred in the extracellular matrix, regulate the diffusion of FGFs [171, 178] and so can determine the shape of FGF concentration gradients in development [179, 180], as well as the storage and release of FGFs in tissue homeostasis [181, 182], The growth factor/morphogen type signals generated by FGFs require the assembly of the ternary complex of FGF ligand, FGF receptor (FGFR) and HS, which engages both the ligand and receptor [118, 125, 164, 165]. Thus, in this respect, HS acts as a co-receptor.

In terms of the specificity of interactions of proteins with HS there are different paradigms and views. One paradigm is the activation of antithrombin III by its binding to a specific pentasaccharide sequence [163], which has been successfully transformed into a synthetic anticoagulant, Arixtra [183, 184]. It was the higher affinity saccharide unit salt-eluted from an antithrombin III affinity column that was originally identified, but it was not the only sequence that bound antithrombin III and able to activate it. More recent findings are that

activity relates to thermal stabilisation of antithrombin III [185] and many oligosaccharide structures have now been shown to possess both a high affinity for antithrombin III and to exert strong anticoagulant activity [186, 187]. With other proteins there is even less consensus. Thus, with FGFs highly specific binding structures in heparin and in HS have been sought [188]. In other experiments, however, the conclusion was that the charge density of the polysaccharide was the major determinant of binding selectivity [162].

In a recent attempt to understand the extent, if any, of selectivity of FGFs for binding to HS, the molecular basis of the interactions between six FGFs from five subfamilies and HS was characterized in depth. This work suggested that there is a degree of selectivity in FGF-heparin interactions and this reflects the evolution of the FGF family members [34], which parallels the specificity of FGF ligands for FGFRs [169].

However, this work is limited in its coverage: two FGFs from one subfamily and one from each of four other subfamilies. Therefore, alternative explanations are quite possible. Consequently, here we characterize the interactions with HS of a further six FGFs from two perspectives. The preference of FGFs for particular sugar structure has been determined using differential scanning fluorimetry (DSF) and a library of chemically modified heparins, heparin derived oligosaccharides and model glycosaminoglycans. A protect and label approach is then used to identify lysine side chains involved in heparin binding and so map the primary and secondary HS binding sites in the FGFs. Pooling the present data with those acquired previously [34] demonstrates that the FGFs show clear selectivity for binding structures and that this, along with the pattern of secondary binding sites on the surface of the FGFs follows the phylogeny established by amino acid sequence alignment. Thus, the molecular basis of the interactions of FGFs with HS, with their preference for particular isoforms of the FGFR has followed the expansion and specialization of the FGF family that occurred during the course of the evolution of the more complex body plans of animals.

Experimental procedures:

Materials Heparin (17 kDa average molecular mass, Celsus Lab, Cincinnati, OH) was used in all assays and as the starting material for the production of modified derivatives and oligosaccharides. Different chemically modified heparin derivatives D1-9 (Table 1; [189]) and cationic forms were produced, as described (Rudd, T.R., 2007) while oligosaccharides with degrees of polymerization (dp) dp 2 to dp 12 were obtained from Iduron (Manchester UK). Porcine mucosal heparan sulfate (HS) and hyaluronic acid (HA) were from Sigma (St. Louis USA); chondroitin sulfate (CS) and dermatan sulfate (DS) were from Iduron.

Recombinant human FGFs

cDNA encoding Histag-FGF-4 (UniProt accession number P08620; residues 31-206) was transformed into C41 (DE3) cells and expressed by inducing with 1 mM isopropyl 1-thio-beta-D-galactopyranoside at 37°C for 3 h. After cell lysis by sonication and clarification by centrifugation, the supernatant was loaded onto an affinity chromatography HiTrap heparin HP column (GE Healthcare, Amersham, Bucks, UK). FGF4 required further purification with a cation-exchange HiTrap 1 ml SP HP column to remove the contaminants, as describe previously (Uniewicz paper). HaloTag (HT)-FGF17 and HaloTag (HT)-FGF6 were expressed and purified as described [171], the HT-FGF17 protein was digested overnight by mixing with TEV protease at ratio 40:1. The sample was then loaded on a HiTrap Q column. FGF17 eluted in the flow through fraction since it did not bind to this anion-exchange matrix, whereas the HaloTag protein bound to the column. The FGF-17 was then further purified by affinity chromatography a Hitrap Heparin column. Histag-FGF-3, HaloTag (HT)-FGF6, Histag-FGF-10, and Histag-FGF-20 were expressed and purified, as described [171].

Size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS)

Analysis of the solution molecular mass were performed by separation of proteins on Superose 200 HR10/300 columns (GE Healthcare) connected in series with a Wyatt Dawn8+ and Wyatt Optilab T-rEX (Wyatt Technology) at 22°C. Samples (100 µL) were filtered and then applied to the column, which was developed in 150 mM NaCl buffer with either 50 mM HEPES or 50 mM Tris-Cl, both pH 7.4 at a flow rate of 0.75 mL/min.

Differential Scanning Fluorimetry (DSF)

DSF was performed with a 7500 Fast Real-Time PCR (RT-PCR) instruction (software version 1.4.0 Applied Biosystems), as described [34, 190]. The different sugars (100 µM, 3.5 µL in HPLC grade water) and FGFs (50 µM, 3.5 µL) as ten-fold concentrated stock solutions, phosphate-buffered saline (PBS, NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 10 mM, KH_2PO_4 1.8 mM; 24.5 µL), and freshly prepared 100-fold stock solution Sypro Orange dye (3.5 µL) (Life Technologies Ltd) were added to a Fast Optical 96 Well Reaction plate (Life Technologies Ltd) kept on ice. After sealing with Optical Adhesive Film (Life Technologies Ltd), the plate was directly analysed in the RT-PCR instrument with a heating cycle covering a gradient between 32 and 81°C in 99 steps of 20 s. First derivatives of the melting curves were calculated with Origin 7 (OriginLab Corp., Northampton, UK). For each sugar, at least two experiments each in triplicate was performed and analysed. The mean melting temperature Tm and the standard error (S.E.) were calculated based on the six repeats. Data were normalized as: [Tm x-Tm PBS] /[Tm hep- Tm PBS], where Tm x is the Tm of protein

in the presence of the heparin derivative; Tm PBS is the Tm of the protein in PBS, and Tm hep is the Tm of the protein in the presence of heparin. The relative stability of protein in PBS buffer was set to 0, while the relative stability of the protein in the presence of heparin was set to 1.

Protect and label identification of lysines involved in heparin binding

Lysine protection: The identification of lysines in heparin binding sites (HBS) was according to Ori *et al*.[170] with minor modifications. A heparin minicolumn was made by placing a plastic air filter at the end of a small pipette tip into which 30 µL of AF-heparin beads (Tosoh Biosciences GmbH, Stuttgart, Germany; binding capacity 4 mg antithrombinIII/mL resin) was packed. A 5 mL syringe was used to pack the minicolumn and dispense buffer. The heparin column was equilibrated 4 times with 50 µL of PB 150 buffer (17.9 mM Na_2HPO_4, 2.1 mM NaH_2PO_4, 150 mM NaCl, pH 7.8). A minimum of 40 µg FGF protein was loaded onto the heparin column and the loading was repeated 3 times with the same sample. After binding, the column was washed with PB 150 buffer 4 times. To acetylate exposed lysines, the minicolumn was then quickly rinsed with 20 µL of PB 150 containing 50 mM sulfo-NHS-acetate (Life Technologies, Paisely UK) and then incubated for 5 min with 20 µL of fresh PB 150 containing 50 mM sulfo-NHS-acetate at room temperature. After acetylation, the minicolumn was washed with 50 µL of PB 150 buffer and acetylated protein was eluted from heparin with 2 x 20 µL elution buffer (45 mM Na_2HPO_4, 5 mM NaH_2PO_4, 2 M NaCl, pH 7.8).

HBS Lysine Biotinylation: Acetylated protein was diluted with 200 µL of PB buffer and concentrated with a 5 kDa MWCO centrifugal filter (Sartourius Ltd., Epsom, UK) by centrifugation for 15 min at 11200 x g. The volume was adjusted to 37.2 µL with PB buffer and any remaining amino groups were biotinylated by the addition of 2.8 µL 145 mM NHS-biotin (Life Technologies, Paisely UK) in dimethylsulfoxide and 30 min incubation at room temperature. The biotinylation reaction was quenched with 4 µL of 1 M Tris, pH 7.5. Then, the sample was transferred to a desalting centrifugal column (7 kDa MWCO, Thermo Scientific, Rockford, UK), covered with 70 µL HPLC grade water and centrifuged for 2 min. Samples were frozen at -80°C for 10 min and dried by centrifugal evaporation.

Protein Digestion: Dried sample was dissolved with 25 µL 8 M urea, 400 mM NH_4HCO_3, pH 7.8 and 2.5 µL 45 mM DTT and incubated for 15 min at 56°C. Then, the samples were carbamidomethylated with 2.5 µL of freshly made 0.1 M iodoacetamide for 15 min at room temperature in the dark. Proteins were diluted with 70 µL of HPLC grade water and digested overnight with 1 µg of MS-grade protease (trypsin, chymotrypsin, thermolysin or Glu-C, Promega UK).

Identification of Labelled Peptides: Biotinylated/acetylated peptides were made up to 0.5 % (w/v) trifluoroacetic acid (TFA) and desalted using C18 Zip Tips (Millipore). The latter were pre-wetted with 100 % (v/v) acetonitrile (ACN) and then pre-equilibrated with 0.1 % (w/v) TFA in water. The peptides were loaded on the Zip Tip and then washed with 10 µL 0.1 % (w/v) TFA. Finally, these peptides were eluted with 2 aliquots of 4-6 µL 50 % (v/v) ACN. The samples were concentrated by rotary evaporation. Analyses were performed on a MALDI-TOF mass spectrometer (Waters, Corp., Manchester, UK). The MS spectra were produced by MassLynx v4.0 and then analysed with the MS-digest tool of the Protein

Prospector package v 5.12.4 with the following parameters: considered modification, acetyl (K), biotin (K), carbamidomethyl (C), carboxymethyl (C); protease used, trypsin/chymotrypsin, thermolysin or Glu-C; missed cleavages, 5; minimum-maximun mass: 800-4000.

Results:

We have used two approaches and multiple representatives of the different human heparin-binding FGF subfamilies, FGF-3 and -10 (FGF-7 subfamily), FGF-16 and -20 (FGF-9 subfamily), FGF-4 and -6 (FGF-4 subfamily), and FGF-8 and -17 (FGF-8 subfamily) to gain an insight into the selectivity and structural basis of the interaction of FGFs with glycosaminoglycans. DSF makes use of a sensitive dye (Sypro Orange), which when bound to aromatic residues produces a high fluorescence; these residues are exposed when proteins are thermally denatured [190]. This allows measurement of the extent of stabilization of the structure of FGFs that occurs upon binding different glycosaminoglycan structures. The protect and label approach identifies lysine residues that are engaged in direct interactions with heparin and so determines the likely binding sites of the polysaccharide in the FGFs. It is capable of identifying lysine residues in both the primary, higher affinity canonical binding site and the much lower affinity secondary binding sites [34, 170].

Thermal stabilization of FGFs by interaction with heparin

The change in fluorescence was measured as temperature was increased with 5 μM of each FGF, and then the first derivative was calculated to determine the melting temperature. Only one melting curve and corresponding derivative per sample is shown in the figures for clarity. Complete data sets are shown in supplementary data. These experiments show that heparin has a concentration-dependent effect on the thermal stability of all the FGFs tested, since their melting temperature progressively increased as the concentration of heparin increased. The melting temperature (Tm) of FGF3 and FGF17 is 36°C and 37.5°C, respectively, while the Tm of FGF4 (49.5°C), FGF10 (41.6°C) and FGF20 (52.2°C) [191, 192] is considerably higher (Figs 1D). In the case of FGF6, the protein aggregates when the N-terminal HaloTag fusion protein is removed [171], so the DSF assay was performed on the fusion protein. Two distinct peaks are observed, one at 46.5°C and the other at 58.5°C (Fig. 5B). The lower melting temperature (46.5°C) is assigned to FGF6 for two reasons. First, the Tm of purified HaloTag corresponds to the second peak (Fig. S3B), whereas the peak at 46.5°C is shifted to higher temperature when the protein is incubated with heparin (Fig. 5A) and only the FGF6 moiety binds the polysaccharide [171]. Thus, not only are FGF4, FGF6, FGF10 and FGF20 more stable than FGF3 and FGF17 in the absence of heparin, but interestingly, human FGF3 and FGF17 would be under these conditions unstable at normal body temperature (Fig. 1D).

To determine the effects of binding heparin, a range of heparin concentrations (0-100 μM) was tested against a fixed concentration (5 μM) of FGFs and the melting temperature was calculated at each concentration of polysaccharide. The melting curves show that a stabilizing effect of heparin on FGF3 and FGF10 is apparent from 0.5 μM heparin (heparin-

FGF molar ratio: 1:10) to 2.5 μM heparin (heparin-FGF molar ratio: 1:2) and is then unchanged at higher concentrations of heparin (Fig. 1C and supplemental Fig. S1C). The melting temperature of FGF6 was increased by heparin though the signal from the more stable HaloTag confounded that from FGF6 when the latter was fully stabilised by heparin. Therefore, this approach could be used to measure the relative stabilizing effect interaction of FGF6 with glycosaminoglycans. The effect of heparin on the melting temperature of FGF4, FGF6 and of FGF17 is similar, with stabilization becoming apparent at 0.5 μM heparin and reaching a maximum around 2.5 μM heparin. However, the stabilizing effect of heparin on FGF20, which is apparent at 0.5 μM heparin does not reach a maximum even with 100 μM heparin, and so is distinct from the other four FGFs. The thermal stabilization of FGF-20 by heparin was the lowest at 10°C, whereas for FGF4, FGF6, FGF10 and FGF17 it ranged from 11°C-15°C and was 17°C for FGF3. Thus, in the case of FGF3 and FGF17, binding to heparin raises their melting temperature well above body temperature.

Analysis of sugar binding selectivity by DSF

The structures in the polysaccharide required for binding these FGFs were then determined by measuring the stabilization effect of a library of model glycosaminoglycan and their derivatives. The molar ratio of FGF: polysaccharide used in this experiment was approximately 1:2 (FGF, 5 μM; polysaccharide, 10 μM).

FGF-3, a member of the FGF7 subfamily according to amino acid sequence alignment was similarly stabilized by unmodified heparin and any of the singly desulfated heparins (D2, D3 and D4) (Fig. 2a). However, the doubly desulfated heparins with just a 6-O-sulfate or N-sulfate (D5 and D7) stabilized FGF-3 to 40 % and 25 % of the level observed with heparin respectively, whereas heparin with just 2-O-sulfate was without a detectable effect. Totally desulfated heparin was also without effect. FGF3 was most stabilized by persulfated heparin, whereas HS was as effective as the doubly desulfated N-sulfated or 6-O sulfated heparins, D5 and D6. FGF-3 did not have a detectable interaction with HA or CS, but DS clearly does bind, albeit not as well as HS. A dp 4 was the shortest oligosaccharide able to stabilize FGF-3 with a maximum effect at dp 10, which stabilized FGF-3 to an extent similar to full-length heparin. FGF3 did not discriminate between the different cationic forms of heparin, since these all had the same stabilizing effect (Fig. 2C). Thus, FGF3 has a clear preference for a saccharide structure with any two of N-, 2-O and 6-O sulfate and is able to bind structures with doubly desulfated heparin containing just a 6-O-sulfate or an N-sulfate, while a dp 10 is likely to represent the full-length binding structure in the polysaccharide.

In the case of FGF10, another member of the FGF7 subfamily, there was no discernible difference in the stabilizing effect of heparin and the singly desulfated heparins (D2, D3 and D4) (Fig. 3a). However, all three doubly desulfated heparins (D5, D6 and D7) had a similar stabilizing effect, which was ~50 % to 70 % of that seen with heparin. Totally desulfated heparin and HA failed to bind FGF10, whereas FGF10 bound persulfated heparin more effectively than heparin and HS only slightly more weakly. FGF10 also bound both CS and DS, though the former more weakly. A heparin derived dp 4 oligosaccharide provided substantial binding and maximum binding was seen with a dp 8, indicating that this is the likely minimum sized fragment of the polysaccharide required for interaction. FGF10 may

also have a slight preference for the Ca^{2+}, Zn^{2+} and Cu^{2+} cation forms of heparin. Thus, the binding preferences of FGF10 are similar, but not identical to those of FGF3. Compared to FGF3, FGF10 has a less marked preference for singly over doubly desulfated heparins, it does not appreciably prefer any of the three doubly desulfated heparins and has a wider range of GAG species (CS as well as DS) with which it can interact (Figs 1B, 1C, 2B, 2C).

For FGF4, there was a greater thermal stabilization by heparin than by any of the singly desulfated heparins (Fig. 4A). Moreover, amongst the singly desulfated heparins, FGF4 had a preference for polysaccharides with both 2-O and N- sulfate (D4) over polysaccharides with N-sulfate and 6-O sulfate (D3), or 2-O and 6-O sulfate (D2). There was no appreciable effect of the doubly desulfated heparins on the thermal stability of FGF-4. These data suggest that the core recognition structure of FGF4 in the polysaccharide involves a 2-O and N-S sulfated structure. The lower stabilization observed with HS may reflect that sequences of the appropriate length containing this motif are relatively rare in this material. FGF4 also bound DS, though weakly, and did not bind to HA or CS (Fig. 4A). FGF4 did not interact detectably with a dp 4 and a dp 6 was the minimal fragment required for binding. Maximal binding, equivalent to that observed with heparin was achieved with a dp 12 (Fig. 4B). Little effect was observed for the different cation coordinated forms of heparin, indicating that this parameter, which changes the conformation of the polysaccharide chain [193], does not, at least in the case of the heparin polysaccharide, influence the binding of FGF4 (Fig. 4C).

FGF6 is another member of the FGF4 subfamily, the singly desulfated heparins stabilized FGF6 more than doubly desulfated heparin. FGF6 showed a strong preference for the combination of 2-O and 6-O sulfated and 2-O and N-sulfated heparin. With the exception of 2-O-sulfated heparin, the doubly desulfated heparins barely stabilized FGF6, which highlights the preference of FGF6 for a structure containing 2-O-sulfate. FGF6 was also stabilized by HS in a similar manner to heparin, whereas HA, CS and DS did not produce any detectable stabilization of FGF6 (Fig. 5C). The minimum size and maximum size of oligosaccharide required for binding to FGF6 was dp 6 and dp 12, respectively (Fig. 5D).

FGF17 bound the three singly desulfated heparins similarly to heparin and the doubly desulfated heparins more weakly (Fig. 6A). Moreover, FGF17 has a mild preference for heparin with either a 2-O or a 6-O sulfate, compared to heparin with just an N-sulfate (Fig. 5A). FGF17 did not bind desulfated heparin or HA, but bound persulfated heparin and HS similarly to heparin. It also interacted with DS to a similar extent as the singly desulfated heparins and more weakly with CS. FGF17 required at least a dp 4 oligosaccharide for binding and maximal binding was observed with a dp 8. It may have a slight preference for Zn^{2+} coordinated heparin over other cationic forms of the polysaccharides (Figs 6B & 6C).

Heparin was more effective at stabilizing FGF20 than any of the singly desulfated heparins (Fig. 7A). FGF20 had a weak interaction with the doubly desulfated heparin possessing just 6-O-sulfate or 2-O-sulfate, whereas there was no detectable interaction with just a N-sulfate, which suggests a preference for the former two sulfation positions. The stabilizing effects of CS and DS on FGF20 were similar to that seen with HS. The minimum size of oligosaccharide required for binding to FGF20 was dp 10 whilst the maximum size of oligosaccharide used in this assay, dp 12, only stabilized the protein to around 40 % of the extent observed with heparin (Fig. 7B). The binding of FGF20 to the polysaccharide was markedly affected by the coordinating cations: the divalent cation (Ca^{2+}, Zn^{2+} or Cu^{2+})

coordinated heparins were twice as effective in stabilizing FGF20 as heparins coordinated to a monovalent cation (Na^{1+} and K^{1+}) (Fig. 7C).

Identification of lysines involved in heparin binding by protect and label

The lysine residues involved in binding heparin in the FGFs were determined by the "protect and label" approach, where lysines in binding sites are protected with acetyl groups whilst the FGFs are bound to heparin [170]. Following release of the FGF from heparin, the newly exposed lysines that had been involved in binding are labelled with biotin and identified by mass spectrometry.

FGF7 Subfamily (FGF3/FGF10)

Initial experiments with FGF3 and FGF10 identified just one peptide with biotinylated lysines, Lys-47 in FGF3 and Lys-81 in FGF10 both in strand β1 (Table 2). This was considered to be due the use of chymotrypsin to cleave the protein, which may produce peptides from these FGFs that are either too long or too short for detection with MALDI-MS. To identify further peptides, FGF3 and FGF10 were digested with trypsin (cleaves at Arg residues only, due to the protect and label procedure) and thermolysin (cleaves at Pro, His, Asp, Glu residues), as modifications of the published method [170]. By changing the protease used to cleave the lysine modified protein, a substantial number of biotinylated lysine residues were identified in FGF3. Lys-160, Lys-168, Lys-174, Lys-204, and Lys-214 were found to be biotinylated. These residues are located the loops between strands β10 and β11 and between strands β11 and β12 and C-terminal to strand β12. They correspond to the HBS1 of FGF-3, predicted by sequence alignment (Fig. 12) [34]. In addition, two other lysine residues, which are close to the above residues of the canonical HBS1 were biotin labelled: Lys-53, which lies between strands β1 and β2, and Lys-101, which is between strands β6 and β7. Thus, these two residues are likely to be part of the canonical binding site, which has contributions from residues that are distant in the primary sequence, but neighbouring in the folded protein. However, biotinylated Lys-47 on strands β1 is distant from the canonical binding site. Along with the neighbouring arginine (Arg-44-46), this would be part of the secondary binding site termed HBS3. The amino acids in FGF3 corresponding to HBS4 identified in FGF7 are arginine and asparagine (Fig. 12, [34]), which would not be detected by the lysine targeted protect and label used here. Thus FGF3 may also possess and HBS4, but this remains to be established.

Lys-87, Lys-184, Lys-191, and Lys-195 were all labelled in FGF10 (Table 2). These lysine residues are in the canonical HBS1 of FGF-10, as predicted by sequence alignment (Fig. 8). Three lysines of the peptide "^{131}YLAMNKKGKLY^{141}" (Lys-125, Lys-126 and Lys-128) of FGF10 were found to be both acetylated and biotinylated (Fig. 8). This has been observed previously [170, 194] and is considered to be due to the local dissociation of a lysine side chain from its interaction with the polysaccharide. In the presence of the NHS-acetate used in the protection step, it becomes acetylated this would likely preclude its re-binding to the polysaccharide. Since the protein remains bound, these lysines must form part of the binding site that is relatively dynamic over the time-scale of the protection step, whereas the remainder of the binding site is not. Two biotinylated lysines were also identified in peptide "IEKNGKVSGTK" (residues, 92-102). Owing to the position of these residues on the

surface of FGF10, they are most likely to form part of HBS4, as identified in FGF7 previously [34] and in FGF3 in the present work (Fig. 8). HBS4 lies orthogonal to HBS1 in all three proteins, thus the binding of the polysacchasride to these sites is mutually exclusive. Similar to FGF3, the biotinylated Lys-81 located in strand β1 has also been identified to be part of HBS3. Since the aligned HBS3 on FGF7 has arginine residues rather than lysines, it was not detected. The identification of HBS3 in FGF3 and FGF10 strongly suggests that the corresponding sequence in FGF7 has the same function.

FGF4 subfamily (FGF4/FGF6)

For FGF4, Lys-183, Lys-186, Lys-188 and Lys-189 in the area between β10 strand and β12 strand were found to be biotinylated (Table 3). These lysine residues correspond to the HBS-1 of FGF4 predicted by sequence alignment (Fig. 9). The mutation to alanine of Lys-183 and Lys-188 in FGF4 had previously identified these residues as being part of the canonical HBS-1 [195]. Moreover, another three labelled lysine residues are physically adjacent: Lys-142 is in the loop between β6 strand and β7 strand; Lys-144 is on the β7 strand; Lys-147 is in the loop between β9 strand and β10 strand. These six residues can be considered to delineate the canonical HBS-1 of FGF4 (Fig. 9). Further biotinylated lysines (Lys-65, Lys-81 and Lys-158) were identified. These three lysines are physically adjacent, but distant from the residues in HBS1. They are aligned with a secondary HBS, HBS3, in other FGFs (Fig. 12) [34, 170]. Similar to FGF4, FGF6 has two heparin binding sites: HBS1, which is identified by biotinylated Lys-144, Lys-185 and Lys-194 (separately, loop between β6 strand and β7 strand, area between β10 strand and β12 strand); HBS3, which includes Lys-83 towards the N-terminal of strand β1 and Lys-158 on the β8 strand (Fig. 9).

FGF8 subfamily (FGF17)

In the case of FGF17, the predicted canonical HBS1 contains arginine, but no lysine residues and, therefore, no peptides were identified between strands β10 and β12. However, Lys-82 and Lys-85, located in the loop between strands β3 and β4, Lys-100 on strand β5, Lys-106 on strand β6 and Lys-119, Lys-123 and Lys-125 on strand β7 and the loop between strands β7 and β8 were all found to be biotinylated. These residues are physically adjacent to the region between strand β10 and β12, where the core of the HBS1 of FGFs is predicted to be located. In addition, FGF17 has an HBS2, which includes Lys-176, Lys-191 and Lys-193 at the C-terminus. The aspartic acid (Asp-121) of FGF18 enlarges the negative border formed by two glutamic acid (Glu-103 and Glu-105), the result of which is that Lys-82 and Lys-100 are part of FGF18's extended HBS2. In contrast, Asp-121 of FGF18 is Ser-121 in FGF17, which consequently has a smaller negatively charged border along its HBS1 and Lys-82 and Lys-100 are now part of an extended HBS1. These data demonstrate how changes in the residues surrounding HBS1 (in this instance Asp to Ser) can alter the structure of an HBS and provide subtle difference between members of the same subfamily (Fig. 10).

FGF9 subfamily (FGF20)

For FGF20, Lys-183, Lys-197, Lys-208 and Lys-212 located in strand β9, the loop between strand β10 and strand β11, and the loop between strand β11 and strand β12 were biotinylated and these correspond to the predicted HBS1 by sequence alignment [34]. Two further lysine residues (Lys-148 and Lys-183) that are located in the area between strand β6 and strand β7 were also identified to be biotinylated. Since these two lysines are physically adjacent to the

canonical binding site, they were considered to be an extension of the HBS1. The biotinylated Lys-231 located in the C-terminus are quite close to Arg-90, Arg-91 and Arg-92, which may form the HBS3 (Fig. 11). However, arginine residues cannot be identified by the NHS chemistry used here. In any event, FGF20 possesses a single, enlarged HBS-1, similar to FGF9 [34], and may also possess an HBS3. The equivalent residue in FGF9, Lys- was not found to be labelled in previous work [34], which may be either due to only one protease being used in this work or to the existence of a very well defined negatively charged border around this residue in FGF20. FGF20 has been found to exist as a non-covalent dimer in solution [192], which is also true for the protein we have produced (Fig. S6). When the enlarged HBS1 is mapped onto the dimer structure, the HBS1 from both FGF20 monomers are joined to form a single large heparin binding surface.

Discussion

To resolve the extent to which molecular specificity underpins the interactions of proteins with HS we have used the phylogenetic relationship of the FGF family, established from amino acid sequence as a test bed. The interactions of FGFs and heparin/HS have been measured from two perspectives: the structures in the polysaccharide required for binding and that lysines in heparin binding sites on the FGFs.

Taking the present results with previous ones [34] provides a comprehensive coverage of the paracrine FGFs, all of which bind HS, with at least two FGFs from each of the subfamilies analysed. Collectively these data follow a clear pattern. FGF members from the same subfamily have a preference for binding similar patterns of sulfation and length of oligosaccharide whereas FGFs from different subfamilies have much more pronounced differences in these preferences (Fig. 13). Moreover, FGFs from the same subfamily possess similar secondary HS binding sites and their primary HBS1 have similar architectures. Again, FGFs from different subfamilies have different combinations of secondary HBSs and their HBS1 differs, particularly with respect to the extent to which amino acids that are distant in sequence, but physically close contribute to the HBS1 (Fig. 13). Thus, in the FGF1 subfamily, both FGF1 and FGF2 have similar preference for N-sulfate and 2-O-sulfate, but FGF1 differs in that is also binds saccharide structures with 6-O-sulfated heparin [190]. This subfamily possesses three heparin binding sites, the primary HBS1 and the secondary binding site HBS2 and 3 [34, 42, 170]. In the FGF4 subfamily, FGF4 prefers structures containing 2-O and N-sulfate, whereas FGF6 binds strongly to 2-O and 6-O-sulfated heparin. Compared to FGF4, FGF6 needs a slightly large structure for minimum binding (dp 6). Both FGF4 and FGF6 have a single secondary binding site, which would correspond to HBS3 in the FGF1 subfamily. In the case of the FGF7 subfamily, FGF7 and FGF10 have preference for a similar sulfate pattern and length of oligosaccharide. However, FGF3 barely binds to doubly desulfated heparin containing only 2-O-sulfate and it required larger structure for full binding. The putative HBS3 of FGF7 and HBS4 of FGF3 contain arginine, but not lysine [34] (Fig. 12), which cannot be identified by our lysine targeted method. The protect and label data, when combined to sequence alignment (Fig. 12) indicate that the FGF7 family possesses two secondary heparin binding sites, HBS3 and HBS4, the latter being physically orthogonal to the canonical HBS1. FGF17 and FGF18, which are in the FGF8 subfamily bind to similar structures containing 6-O-sulfate and N-sulfate and they contain a single

secondary binding site, HBS2. In the FGF9 subfamily, FGF9 and FGF20 show similar preference to 6-O-sulfate heparin. FGF9 prefers to bind with N-sulfate rather than 2-O-sulfated heparin, in contrast, FGF20 binds strongly to 2-O-sulfated heparin. Although FGF20 required larger structures for binding than FGF9, this could be caused by its dimer structure. Both FGF9 and FGF20 possess a single, enlarged HBS1. A secondary HBS3 was also been found in FGF20; whereas corresponding basic residues are present in FGF9, in the latter there is no negatively charged border and so they are more likely to be part of an extended HBS1. In contrast in FGF20, these residues are surrounded by a negative border, which would isolate them. Consequently, FGF20 seems likely to have a distinct HBS3.

The expansion of the FGF family and its divergence into subfamilies occurred through genome duplication events that led to even more complex animal body plans [33]. It is clear that the molecular specificity of the FGFs for particular structures in HS and the pattern of secondary binding sites on the FGFs also underwent the similar diversification. This implies that the differences we observe (Fig. 13) are linked to the functional differences that exist between FGF subfamilies. It is important to note that the molecular specificity is far from absolute. That is, there is a consensus ranking of sulfations, oligosaccharide length and GAG preference for FGFs in the same subfamily. This is similar to what is seen with respect to the specificity of FGFs for their receptors [35, 51] and borne out by an analysis of nine FGFs [169]. Indeed, it is possible to identify excellent binding structures in sulfated polysaccharides that are unrelated to GAGs [196]. This indicates that it is the spatial disposition of sulfate, carboxyl and hydroxyl groups on the polysaccharide that are important for binding. This also is likely to underlie the observation that, in addition to its classic pentasaccharide binding sequence [174], which has been the underpinning of arguments for absolute specificity of protein-HS interactions, there are good antithrombin III binding structures with anticoagulant activity that are substantially different [186, 197]. The evolutionary divergence and, within FGF subfamilies conservation of HS binding properties, indicates that they are likely to have functional importance. This is suggested by the measurement of the diffusion of FGFs in pericellular matrix [171], which shows that the diffusion properties of FGF are determined at least in part by their binding specificities for GAGs and by the crosslinking of HS chains by FGF2, but not by FGF9, which has just a single HBS [198].

References

1 Ori, A., Wilkinson, M. C., Fernig, D. G. 2008 The heparanome and regulation of cell function: structures, functions and challenges. *Frontiers in bioscience : a journal and virtual library*. **13**, 4309-4338.

2 Ori, A., Wilkinson, M. C., Fernig, D. G. 2011 A systems biology approach for the investigation of the heparin/heparan sulfate interactome. *The Journal of biological chemistry*. **286**, 19892-19904. (10.1074/jbc.M111.228114)

3 Xu, D., Esko, J. D. 2014 Demystifying heparan sulfate-protein interactions. *Annual review of biochemistry*. **83**, 129-157. (10.1146/annurev-biochem-060713-035314)

4 Gallagher, J. 2015 Fell-Muir Lecture: Heparan sulphate and the art of cell regulation: a polymer chain conducts the protein orchestra. *International journal of experimental pathology*. **96**, 203-231. (10.1111/iep.12135)

5 Nunes, Q. 2015 The role of Heparin-binding proteins in normal pancreas and acute pancreatitis: University of Liverpool.

6 Damus, P. S., Hicks, M., Rosenberg, R. D. 1973 Anticoagulant action of heparin. *Nature*. **246**, 355-357.

7 Rapraeger, A. C., Krufka, A., Olwin, B. B. 1991 Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science*. **252**, 1705-1708.

8 Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., Ornitz, D. M. 1991 Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell*. **64**, 841-848.

9 Dreyfuss, J. L., Regatieri, C. V., Jarrouge, T. R., Cavalheiro, R. P., Sampaio, L. O., Nader, H. B. 2009 Heparan sulfate proteoglycans: structure, protein interactions and cell signaling. *Anais da Academia Brasileira de Ciencias*. **81**, 409-429.

10 Tumova, S., Woods, A., Couchman, J. R. 2000 Heparan sulfate proteoglycans on the cell surface: versatile coordinators of cellular functions. *The international journal of biochemistry & cell biology*. **32**, 269-288.

11 Lin, X. 2004 Functions of heparan sulfate proteoglycans in cell signaling during development. *Development*. **131**, 6009-6021. (10.1242/dev.01522)

12 Hagner-Mcwhirter, A., Lindahl, U., Li, J. 2000 Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. *The Biochemical journal*. **347 Pt 1**, 69-75.

13 Rosenberg, R. D., Shworak, N. W., Liu, J., Schwartz, J. J., Zhang, L. 1997 Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *The Journal of clinical investigation*. **100**, S67-75.

14 Kusche-Gullberg, M., Kjellen, L. 2003 Sulfotransferases in glycosaminoglycan biosynthesis. *Current opinion in structural biology*. **13**, 605-611.

15 Connell, B. J., Lortat-Jacob, H. 2013 Human immunodeficiency virus and heparan sulfate: from attachment to entry inhibition. *Frontiers in immunology*. **4**, 385. (10.3389/fimmu.2013.00385)

16 Itoh, N. 2007 The Fgf families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease. *Biological & pharmaceutical bulletin*. **30**, 1819-1825.

17 Itoh, N., Ornitz, D. M. 2004 Evolution of the Fgf and Fgfr gene families. *Trends in genetics : TIG*. **20**, 563-569. (10.1016/j.tig.2004.08.007)

18 Sun, C., Li, Y., Taylor, S. E., Mao, X., Wilkinson, M. C., Fernig, D. G. 2015 HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. *PeerJ*. **3**, e1060. (10.7717/peerj.1060)

19 Duchesne, L., Octeau, V., Bearon, R. N., Beckett, A., Prior, I. A., Lounis, B., Fernig, D. G. 2012 Transport of fibroblast growth factor 2 in the pericellular matrix is controlled by the spatial distribution of its binding sites in heparan sulfate. *PLoS biology*. **10**, e1001361. (10.1371/journal.pbio.1001361)

20 Vemaraju, S., Kantarci, H., Padanad, M. S., Riley, B. B. 2012 A spatial and temporal gradient of Fgf differentially regulates distinct stages of neural development in the zebrafish inner ear. *PLoS genetics*. **8**, e1003068. (10.1371/journal.pgen.1003068)
21 Lahti, L., Saarimaki-Vire, J., Rita, H., Partanen, J. 2011 FGF signaling gradient maintains symmetrical proliferative divisions of midbrain neuronal progenitors. *Developmental biology*. **349**, 270-282. (10.1016/j.ydbio.2010.11.008)
22 Bashkin, P., Doctrow, S., Klagsbrun, M., Svahn, C. M., Folkman, J., Vlodavsky, I. 1989 Basic Fibroblast Growth-Factor Binds to Subendothelial Extracellular-Matrix and Is Released by Heparitinase and Heparin-Like Molecules. *Biochemistry*. **28**, 1737-1743. (Doi 10.1021/Bi00430a047)
23 Vlodavsky, I., Folkman, J., Sullivan, R., Fridman, R., Ishaimichaeli, R., Sasse, J., Klagsbrun, M. 1987 Endothelial Cell-Derived Basic Fibroblast Growth-Factor - Synthesis and Deposition into Subendothelial Extracellular-Matrix. *Proceedings of the National Academy of Sciences of the United States of America*. **84**, 2292-2296. (DOI 10.1073/pnas.84.8.2292)
24 Sarrazin, S., Lamanna, W. C., Esko, J. D. 2011 Heparan sulfate proteoglycans. *Cold Spring Harbor perspectives in biology*. **3**, (10.1101/cshperspect.a004952)
25 Guimond, S., Maccarana, M., Olwin, B. B., Lindahl, U., Rapraeger, A. C. 1993 Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4. *The Journal of biological chemistry*. **268**, 23906-23914.
26 Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E., Leder, P. 1992 Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Molecular and cellular biology*. **12**, 240-247.
27 Yayon, A., Aviezer, D., Safran, M., Gross, J. L., Heldman, Y., Cabilly, S., Givol, D., Katchalski-Katzir, E. 1993 Isolation of peptides that inhibit binding of basic fibroblast growth factor to its receptor from a random phage-epitope library. *Proceedings of the National Academy of Sciences of the United States of America*. **90**, 10643-10647.
28 Thunberg, L., Backstrom, G., Lindahl, U. 1982 Further characterization of the antithrombin-binding sequence in heparin. *Carbohydrate research*. **100**, 393-410.
29 Chen, J., Avci, F. Y., Munoz, E. M., McDowell, L. M., Chen, M., Pedersen, L. C., Zhang, L., Linhardt, R. J., Liu, J. 2005 Enzymatic redesigning of biologically active heparan sulfate. *The Journal of biological chemistry*. **280**, 42817-42825. (10.1074/jbc.M504338200)
30 Petitou, M., van Boeckel, C. A. 2004 A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? *Angewandte Chemie*. **43**, 3118-3133. (10.1002/anie.200300640)
31 Lima, M. A., Hughes, A. J., Veraldi, N., Rudd, T. R., Hussain, R., Brito, A. S., Chavante, S. F., Tersariol, I. I., Siligardi, G., Nader, H. B., *et al.* 2013 Antithrombin stabilisation by sulfated carbohydrates correlates with anticoagulant activity. *MedChemComm*. **4**, 870-873. (10.1039/c3md00048f)
32 Petitou, M., Duchaussoy, P., Driguez, P. A., Herault, J. P., Lormeau, J. C., Herbert, J. M. 1999 New synthetic heparin mimetics able to inhibit thrombin and factor Xa. *Bioorganic & medicinal chemistry letters*. **9**, 1155-1160.
33 Petitou, M., Herault, J. P., Bernat, A., Driguez, P. A., Duchaussoy, P., Lormeau, J. C., Herbert, J. M. 1999 Synthesis of thrombin-inhibiting heparin mimetics without side effects. *Nature*. **398**, 417-422. (10.1038/18877)
34 Wang, F., Kan, M., McKeehan, K., Jang, J. H., Feng, S., McKeehan, W. L. 1997 A homeo-interaction sequence in the ectodomain of the fibroblast growth factor receptor. *The Journal of biological chemistry*. **272**, 23887-23895.
35 Kreuger, J., Spillmann, D., Li, J. P., Lindahl, U. 2006 Interactions between heparan sulfate and proteins: the concept of specificity. *The Journal of cell biology*. **174**, 323-327. (10.1083/jcb.200604035)

36 Xu, R., Ori, A., Rudd, T. R., Uniewicz, K. A., Ahmed, Y. A., Guimond, S. E., Skidmore, M. A., Siligardi, G., Yates, E. A., Fernig, D. G. 2012 Diversification of the structural determinants of fibroblast growth factor-heparin interactions: implications for binding specificity. *The Journal of biological chemistry*. **287**, 40061-40073. (10.1074/jbc.M112.398826)
37 Xu, R., Rudd, T. R., Hughes, A. J., Siligardi, G., Fernig, D. G., Yates, E. A. 2013 Analysis of the fibroblast growth factor receptor (FGFR) signalling network with heparin as coreceptor: evidence for the expansion of the core FGFR signalling network. *The FEBS journal*. **280**, 2260-2270. (10.1111/febs.12201)
38 Yates, E. A., Santini, F., Guerrini, M., Naggi, A., Torri, G., Casu, B. 1996 1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydrate research*. **294**, 15-27.
39 Uniewicz, K. A., Ori, A., Xu, R., Ahmed, Y., Wilkinson, M. C., Fernig, D. G., Yates, E. A. 2010 Differential scanning fluorimetry measurement of protein stability changes upon binding to glycosaminoglycans: a screening test for binding specificity. *Analytical chemistry*. **82**, 3796-3802. (10.1021/ac100188x)
40 Ori, A., Free, P., Courty, J., Wilkinson, M. C., Fernig, D. G. 2009 Identification of heparin-binding sites in proteins by selective labeling. *Molecular & cellular proteomics : MCP*. **8**, 2256-2265. (10.1074/mcp.M900031-MCP200)
41 Fan, H., Vitharana, S. N., Chen, T., O'Keefe, D., Middaugh, C. R. 2007 Effects of pH and polyanions on the thermal stability of fibroblast growth factor 20. *Molecular pharmaceutics*. **4**, 232-240. (10.1021/mp060097h)
42 Maity, H., Karkaria, C., Davagnino, J. 2009 Effects of pH and arginine on the solubility and stability of a therapeutic protein (Fibroblast Growth Factor 20): relationship between solubility and stability. *Current pharmaceutical biotechnology*. **10**, 609-625.
43 Chevalier, F., Lucas, R., Angulo, J., Martin-Lomas, M., Nieto, P. M. 2004 The heparin-Ca(2+) interaction: the influence of the O-sulfation pattern on binding. *Carbohydrate research*. **339**, 975-983. (10.1016/j.carres.2003.12.023)
44 Uniewicz, K. A., Ori, A., Ahmed, Y. A., Yates, E. A., Fernig, D. G. 2014 Characterisation of the interaction of neuropilin-1 with heparin and a heparan sulfate mimetic library of heparin-derived sugars. *PeerJ*. **2**, e461. (10.7717/peerj.461)
45 Bellosta, P., Iwahori, A., Plotnikov, A. N., Eliseenkova, A. V., Basilico, C., Mohammadi, M. 2001 Identification of receptor and heparin binding sites in fibroblast growth factor 4 by structure-based mutagenesis. *Molecular and cellular biology*. **21**, 5946-5957.
46 Baird, A., Schubert, D., Ling, N., Guillemin, R. 1988 Receptor- and heparin-binding domains of basic fibroblast growth factor. *Proceedings of the National Academy of Sciences of the United States of America*. **85**, 2324-2328.
47 Itoh, N., Ornitz, D. M. 2008 Functional evolutionary history of the mouse Fgf gene family. *Developmental dynamics : an official publication of the American Association of Anatomists*. **237**, 18-27. (10.1002/dvdy.21388)
48 Zhang, X., Ibrahimi, O. A., Olsen, S. K., Umemori, H., Mohammadi, M., Ornitz, D. M. 2006 Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. *The Journal of biological chemistry*. **281**, 15694-15700. (10.1074/jbc.M601252200)
49 Ornitz, D. M., Xu, J., Colvin, J. S., McEwen, D. G., MacArthur, C. A., Coulier, F., Gao, G., Goldfarb, M. 1996 Receptor specificity of the fibroblast growth factor family. *The Journal of biological chemistry*. **271**, 15292-15297.
50 Rudd, T. R., Uniewicz, K. A., Ori, A., Guimond, S. E., Skidmore, M. A., Gaudesi, D., Xu, R., Turnbull, J. E., Guerrini, M., Torri, G., *et al.* 2010 Comparable stabilisation, structural changes and activities can be induced in FGF by a variety of HS and non-GAG analogues: implications for sequence-activity relationships. *Organic & biomolecular chemistry*. **8**, 5390-5397. (10.1039/c0ob00246a)

51 Petitou, M., Herault, L. P., Bernat, A., Driguez, P. A., Duchaussoy, P., Lormeau, J. C., Herbert, J. M. 1999 Synthesis of thrombin-inhibiting heparin mimetics without side effects. *Nature*. **398**, 417-422.
52 Migliorini, E., Thakar, D., Kuhnle, J., Sadir, R., Dyer, D. P., Li, Y., Sun, C., Volkman, B. F., Handel, T. M., Coche-Guerente, L., *et al.* 2015 Cytokines and growth factors cross-link heparan sulfate. *Open biology*. **5**, (10.1098/rsob.150046)
53 Yeh, B. K., Igarashi, M., Eliseenkova, A. V., Plotnikov, A. N., Sher, I., Ron, D., Aaronson, S. A., Mohammadi, M. 2003 Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 2266-2271. (10.1073/pnas.0436500100)
54 Bellosta, P., Iwahori, A., Plotnikov, A. N., Eliseenkova, A. V., Basilico, C., Mohammadi, M. 2001 Identification of receptor and heparin binding sites in fibroblast growth factor 4 by structure-based mutagenesis. *Molecular and cellular biology*. **21**, 5946-5957. (Doi 10.1128/Mcb.21.17.5946-5957.2001)
55 Olsen, S. K., Li, J. Y. H., Bromleigh, C., Eliseenkova, A. V., Ibrahimi, O. A., Lao, Z. M., Zhang, F. M., Linhardt, R. J., Joyner, A. L., Mohammadi, M. 2006 Structural basis by which alternative splicing modulates the organizer activity of FGF8 in the brain. *Gene Dev*. **20**, 185-198. (10.1101/gad.1365406)
56 Kalinina, J., Byron, S. A., Makarenkova, H. P., Olsen, S. K., Eliseenkova, A. V., Larochelle, W. J., Dhanabal, M., Blais, S., Ornitz, D. M., Day, L. A., *et al.* 2009 Homodimerization Controls the Fibroblast Growth Factor 9 Subfamily's Receptor Binding and Heparan Sulfate-Dependent Diffusion in the Extracellular Matrix. *Molecular and cellular biology*. **29**, 4663-4678. (10.1128/Mcb.01780-08)
57 Huson, D. H., Scornavacca, C. 2012 Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. *Syst Biol*. **61**, 1061-1067. (10.1093/sysbio/sys062)
58 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., *et al.* 2007 Clustal W and clustal X version 2.0. *Bioinformatics*. **23**, 2947-2948. (10.1093/bioinformatics/btm404)

Table 1 Nomenclature and structures of chemically modified heparin structures.

Analogue	Predominant repeat	IdoUA-2	GlcN-6	GlcN-2	IdoUA-3	GlcN-3
Heparin	$I_{2S}A^{6S}{}_{NS}$	SO_3^-	SO_3^-	SO_3^-	OH	OH
2	$I_{2S}A^{6S}{}_{NAC}$	SO_3^-	SO_3^-	$COCH_3$	OH	OH
3	$I_{2OH}A^{6S}{}_{NS}$	OH	SO_3^-	SO_3^-	OH	OH
4	$I_{2S}A^{6OH}{}_{NS}$	SO_3^-	OH	SO_3^-	OH	OH
5	$I_{2OH}A^{6S}{}_{NAC}$	OH	SO_3^-	$COCH_3$	OH	OH
6	$I_{2S}A^{6OH}{}_{NAC}$	SO_3^-	OH	$COCH_3$	OH	OH
7	$I_{2OH}A^{6OH}{}_{NS}$	OH	OH	SO_3^-	OH	OH
8	$I_{2OH}A^{6OH}{}_{NAC}$	OH	OH	$COCH_3$	OH	OH
9	$I_{2S,3S}A^{6S}{}_{3S,NS}$	SO_3^-	SO_3^-	SO_3^-	SO_3^-	SO_3^-

Letter I stands for iduronate, and A stands for the amino sugar glucosamine. [a] Numbers refer to the ring position of carbon atoms.

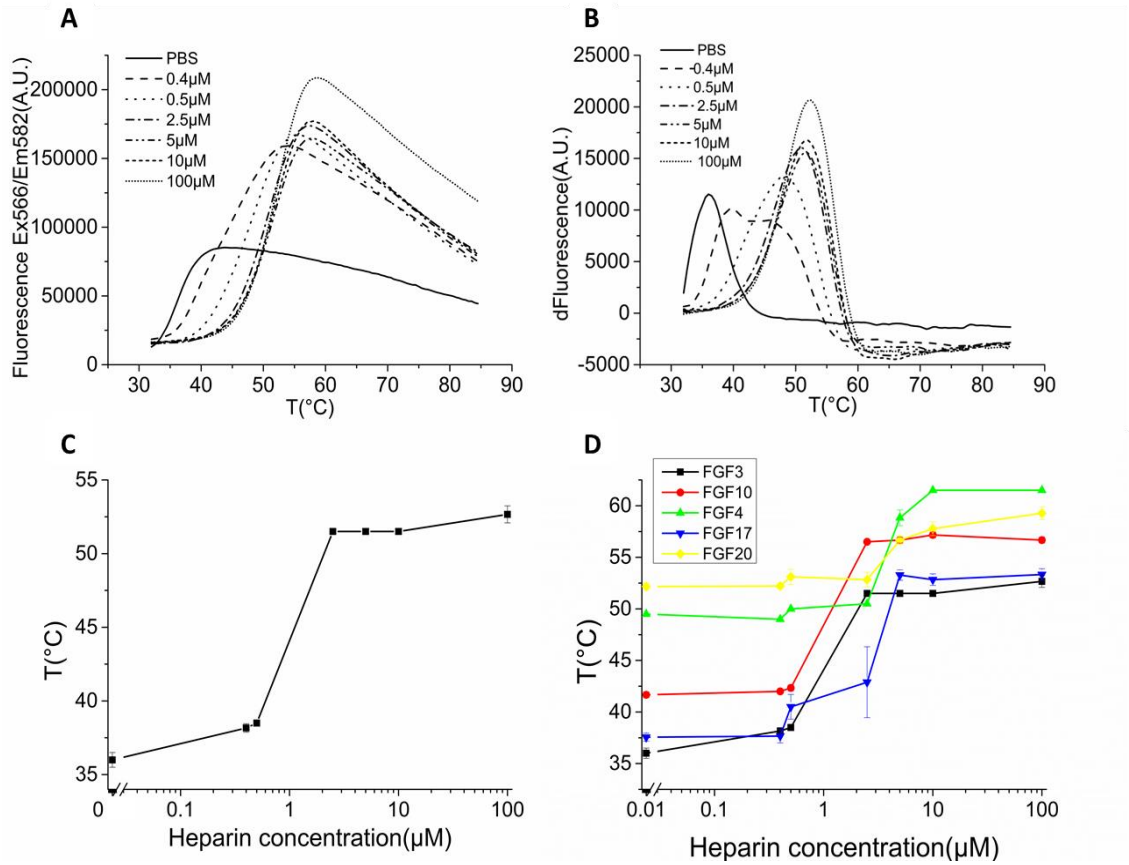


Figure 1 Stabilization effect of heparin on FGF3. Differential scanning fluorimetry of 5 μM FGF3 in the presence of varying concentrations of heparin. **A**, melting curve profiles of FGF-3 (5 μM) with a range of heparin concentrations (0-100 μM). **B**, first derivative of the melting curves of FGF-3 in A. **C**, peak of the first derivative of the melting curves from B, which is the melting temperature, Tm (mean of triplicate ±S.E.). **D**, the curve of melting temperature of five FGFs (FGF3, FGF10, FGF4, FGF17 and FGF20).

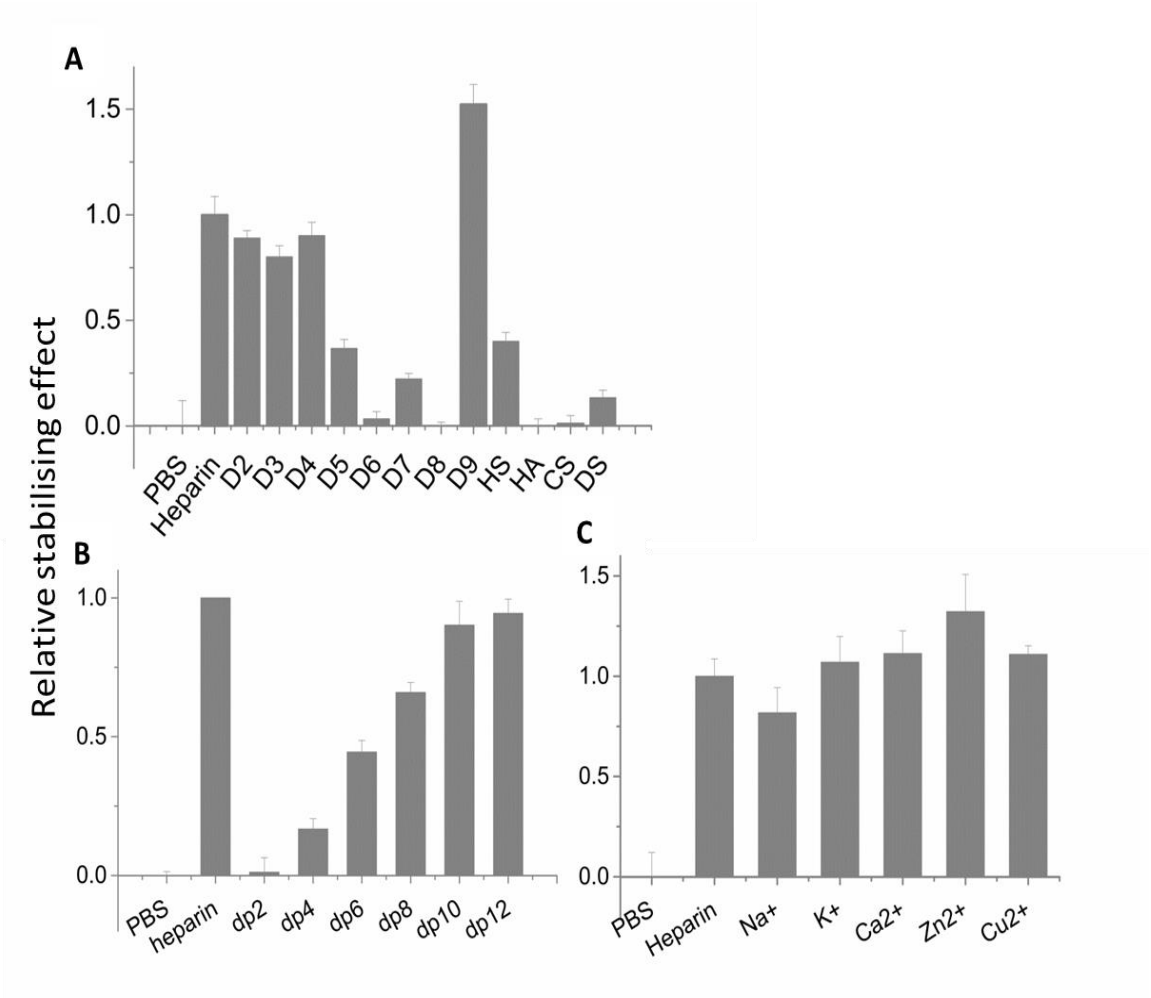


Figure 2 Differential scanning fluorimetry analysis of binding of GAG derivatives to FGF3. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS. **B**, heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **C**, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

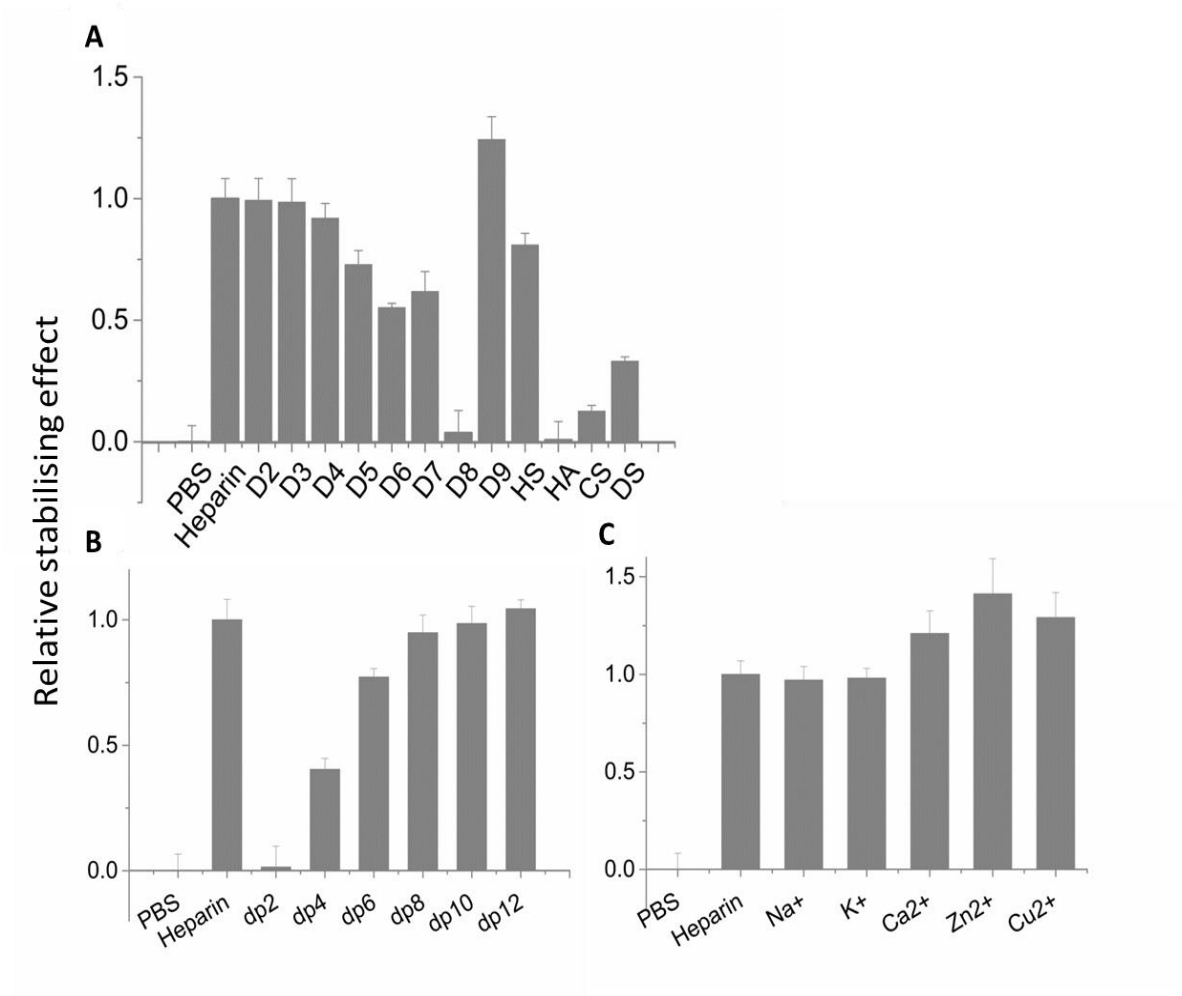


Figure 3 Differential scanning fluorimetry analysis of binding of GAG derivatives to FGF10. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS). **B**, heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **C**, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

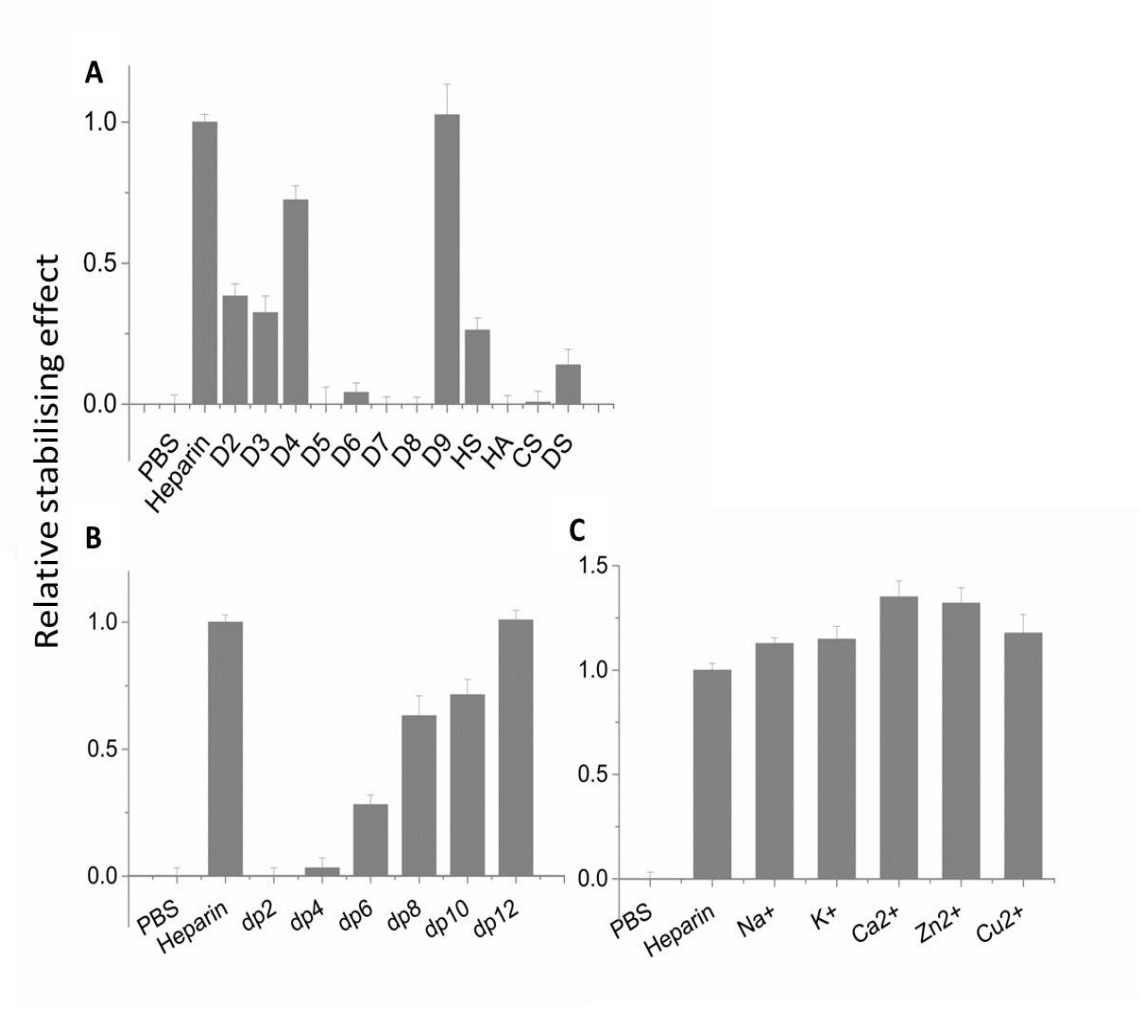


Figure 4 Differential scanning fluorimetry analysis of binding of GAG derivatives to FGF4. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS). **B**, heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **C**, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

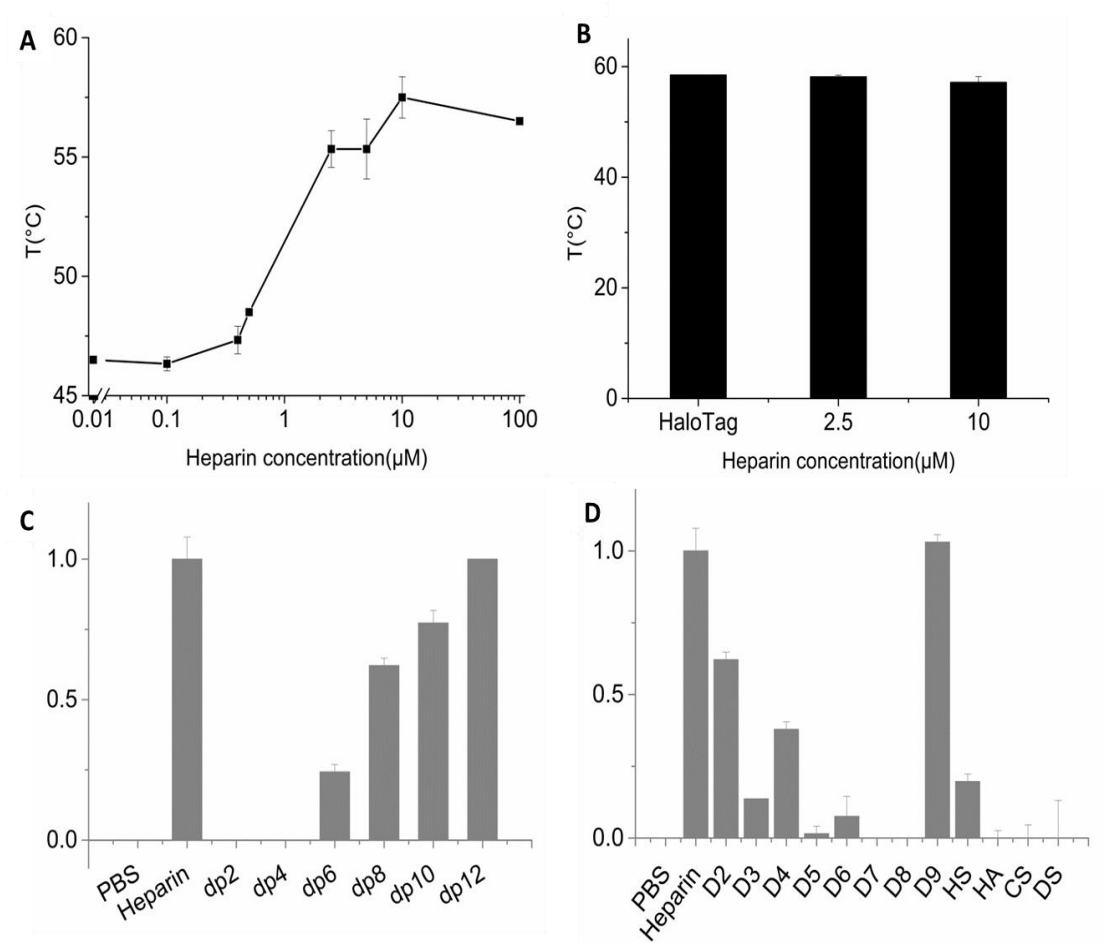


Figure 5 Differential scanning fluorimetry analysis of binding of GAG derivatives to HTFGF6. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, the melting temperature of FGF10 in the presence of varying concentrations of heparin, Tm (mean of triplicate S.E.). **B**, the melting temperature of Halo protein with and without heparin (2.5 μM and 10 μM). **C**, Halotag heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **D**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS). Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

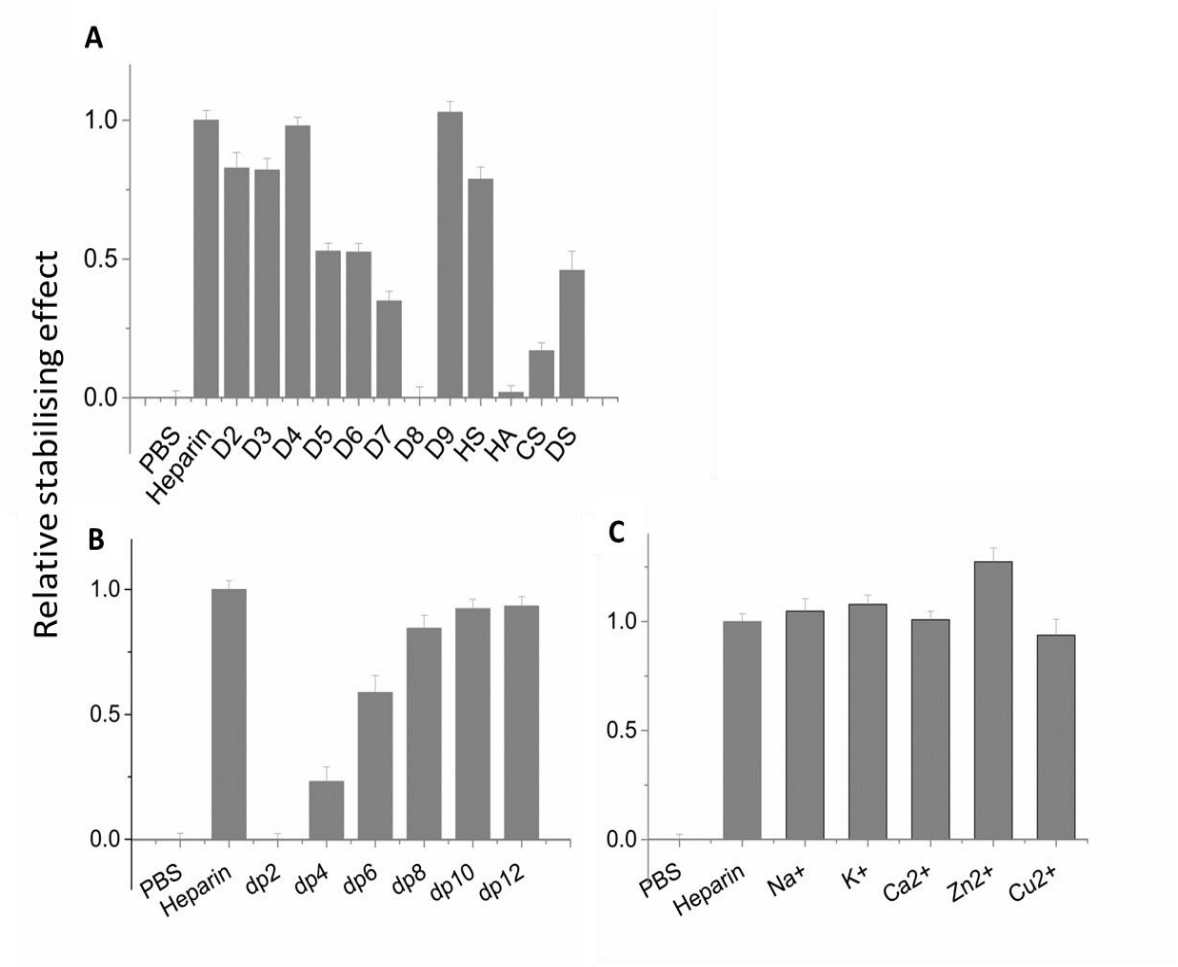


Figure 6 Differential scanning fluorimetry analysis of binding of GAG derivatives on FGF17. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS). **B**, heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **C**, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

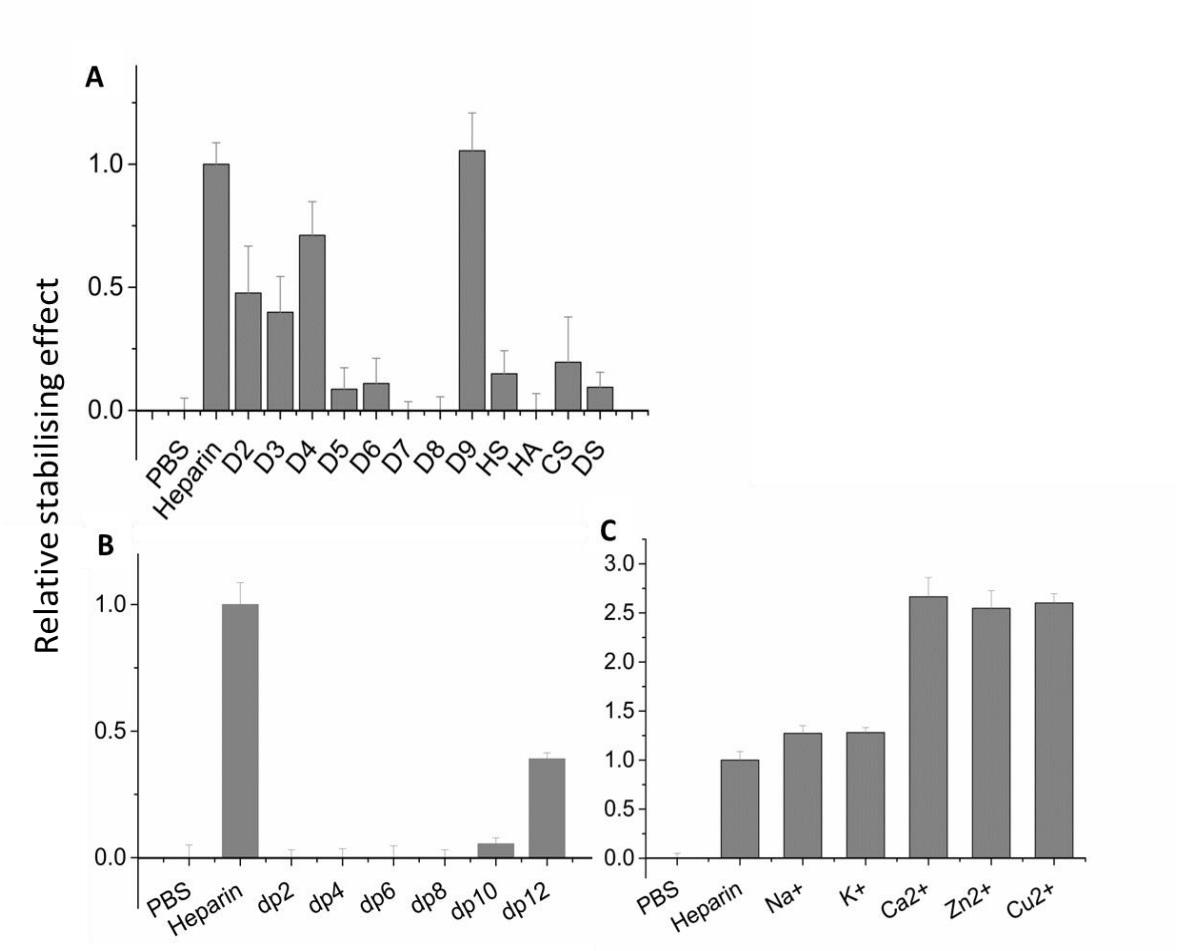


Figure 7 Differential scanning fluorimetry analysis of binding of GAG derivatives on FGF20. Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 10 μM sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated (see "Experimental Procedures"). **A**, thermal stabilization effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS). **B**, heparin-derived oligosaccharides, ranging from dp 2 to dp 12. **C**, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of error bar is due to a small S.E.

Table 2

Summary of peptides of FGF3 and FGF10 identified by lysine targeted Protect and Label structural proteomics

Labelled peptides were identified by MALDI-Q-TOF and analysed by MS-digest of the package of ProteinProspector v 5.12.3. A full list of identified peptides is provided in the supplemental table. The four proteases used for protein digestion were trypsin (TRY), thermolysin (THE), chymotrypsin (CHY) and Glu-C (GLU).

Peptide		Sequences	Proteases	Residues	HBS
FGF3	1	ATK(biotin)YHLQ	(THE)	51-57	1
	2	AMNK(biotin)RGR	(THE)	98-104	1
	3	LWYVSVNGK(biotin)GRPR	(TRY)	152-164	1
	4	RGFK(biotin)TR	(TRY)	165-170	1
	5	TQK(biotin)SSLFLPR	(TRY)	172-181	1
	6	QLQSGLPRPPGK(biotin)GVQPR	(TRY)	193-209	1
	7	QK(biotin)QSPDNLEPSHVQASR	(TRY)	213-229	1
	8	EHLGGAPRRRK(biotin)L	(CHY)	37-48	4
FGF10	1	QMYVALNGK(biotin)GAPR	(TRY)	175-187	1
	2	RGQK(biotin)TR	(TRY)	188-193	1
	3	K(biotin)NTSAHFLPMVVHS	(TRY)	195-208	1
	4	RK(biotin)LFSFTK(biotin)Y	(CHY)	80-88	1/3
	5	IEKNGKVSGTK (2xbiotin)	(TRY)	92-102	4
	6	YLAMNK(biotin/acetyl)K(biotin/acetyl)GK(biotin/acetyl)LY	(CHY)	131-141	4

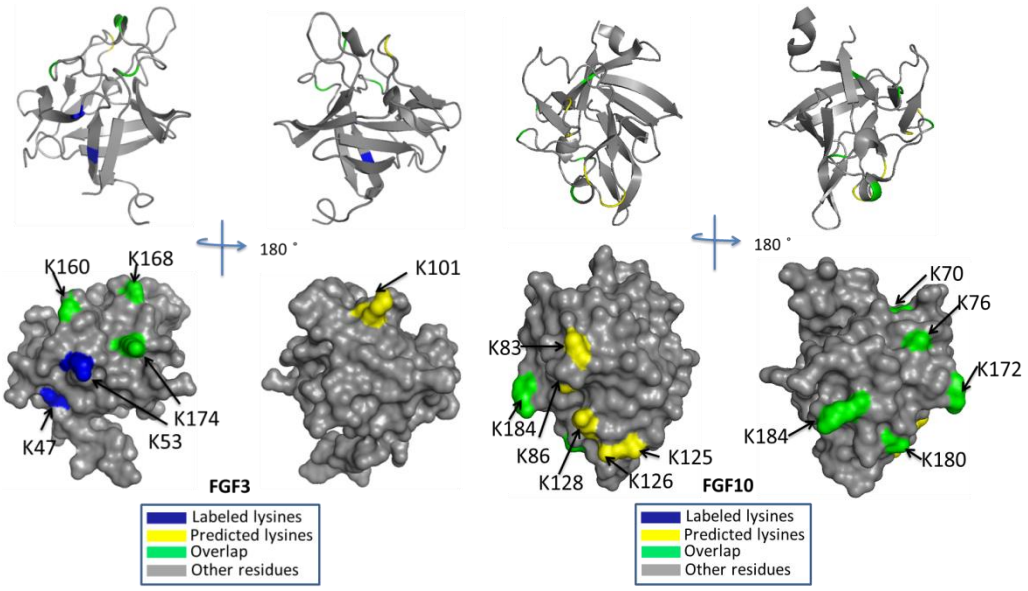


Figure 8 Position of biotinylated peptides in FGF-3 (residues 56-202) identified by structural proteomics mapped onto the predicted three dimensional structure（PDB 1NUN) and FGF-10 (69-207) from（PDB 1NUN) [199]. Protect and label was used to identify the heparin binding sides of FGFs. Labelled peptides are coloured in *blue* and predicted lysines are coloured in *yellow*. The peptides overlapping with aligned HBS lysines are coloured in *green.*

Table 3

Summary of peptides of FGF4 and FGF6 identified by lysine targeted Protect and Label structural proteomics

Labelled peptides were identified by MALDI-Q-TOF and analysed by MS-digest of the package of ProteinProspector v 5.12.3. A full list of identified peptides is provided in the supplemental table. The four proteases used for protein digestion were trypsin (TRY), thermolysin (THE), chymotrypsin (CHY), Glu-C (GLU).

Peptide		Sequences	Proteases	Residues	HBS
FGF4	1	VAMSSK(biotin)GK(biotin)LY	(CHY)	137-146	1
	2	LPNNYNAYESYK(biotin)YPGMF	(CHY)	162-178	1
	3	LSK(biotin)NGK(biotin)TK(biotin)K(biotin)GNRVSPT	(THE)	181-196	1
	4	AQPK(biotin)EAAVQSGAGDY	(THE)	62-76	3
	5	LLGIK(biotin)RL	(CHY)	77-83	3
	6	FK(biotin)EILLPNNYN	(THE)	157-167	3
FGF6	1	SALFVAMNSK(biotin)GR	(TRY)	135-146	1
	2	IALSK(biotin)Y	(CHY)	181-186	1
	3	GSK(biotin)VSPIMTVTHFLPR	(TRY)	192-207	1
	4	SRAGLAGEIAGVNWESGYLVGIK(biotin)RQRR	(TRY)	61-87	3
	5	LYATPSFQEEC(carbamidomethyl)K(biotin)FR	(TRY)	147-160	3

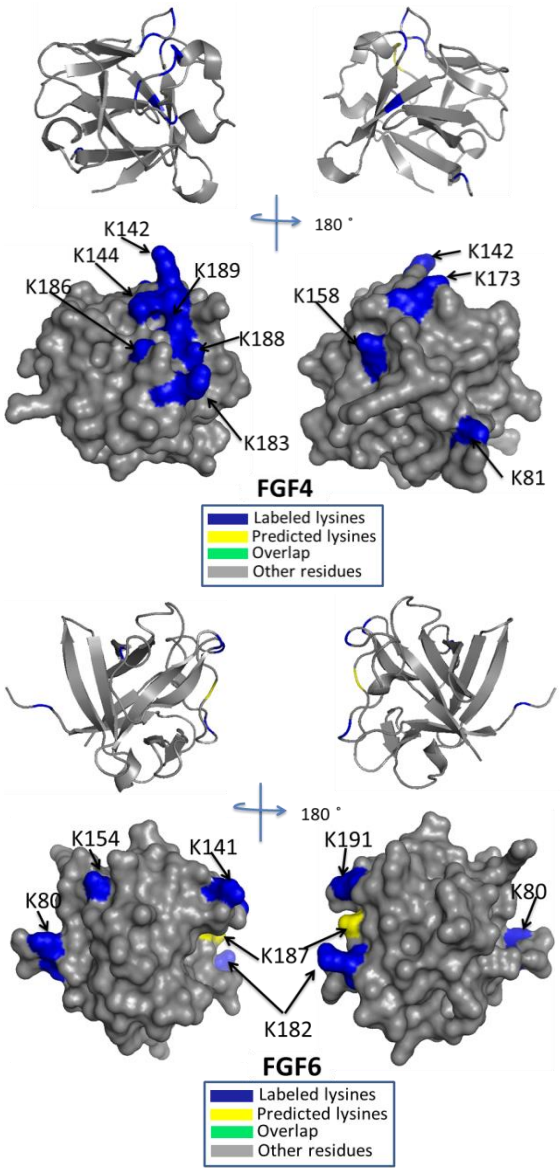


Figure 9 Position of biotinylated peptides in FGF-4 (residues 79-206) and FGF-6 (residues 47-174) identified by structural proteomics mapped onto the predicted three dimensional structure（PDB 1IJT) [200]. Protect and label was used to identify the heparin binding sides of FGFs. Labelled peptides are coloured in *blue* and predicted lysines are coloured in *yellow*. The peptides overlapping with aligned HBS lysines are coloured in *green.*

Table 4

Summary of peptides of FGF17 identified by lysine targeted Protect and Label structural proteomics

Labelled peptides were identified by MALDI-Q-TOF and analysed by MS-digest of the package of ProteinProspector v 5.12.3. A full list of identified peptides is provided in the supplemental table. The four proteases used for protein digestion were trypsin (TRY), thermolysin (THE), chymotrypsin (CHY), Glu-c (GLU).

Peptide		Sequences	Proteases	Residues	HBS
FGF17	1	GNK(biotin)FAK(biotin)LIVETD	(GLU)	80-91	1
	2	GSRVRIK(biotin)GAESEK(biotin)Y	(CHY)	94-107	1
	3	LIGK(biotin)PSGK(biotin)SK(biotin)DCVFTE	(THE)	116-131	1
	4	IK (biotin)RLY	(CHY)	175-179	2
	5	QGQLPFPNHAEK(biotin)QK(biotin)QF	(CHY)	180-195	2

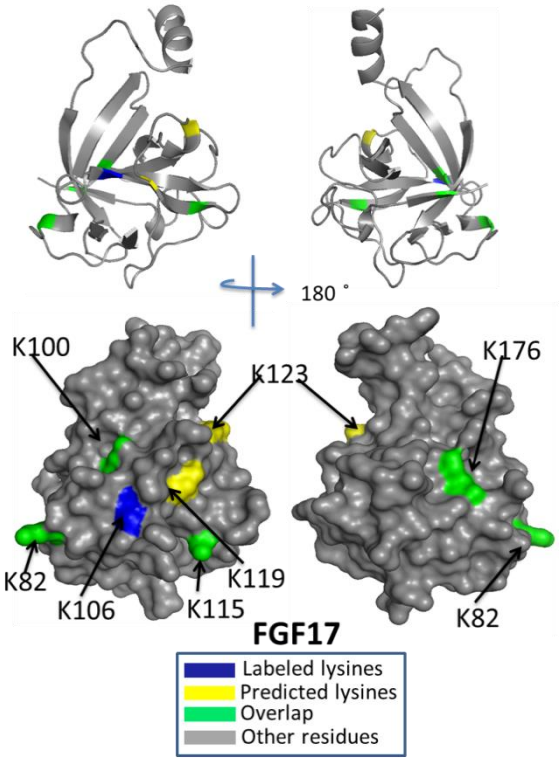


Figure 10 Position of biotinylated peptides in FGF-17 (residues 33-178) identified by structural proteomics mapped onto the predicted three dimensional structure（PDB 2FDB) [201]. Protect and label was used to identify the heparin binding sides of FGFs. Labelled peptides are coloured in *blue* and predicted lysines are coloured in *yellow*. The peptides overlapping with aligned HBS lysines are coloured in *green.*

Table 5

Summary of peptides of FGF20 identified by lysine targeted Protect and Label structural proteomics

Labelled peptides were identified by MALDI-Q-TOF and analysed by MS-digest of the package of ProteinProspector v 5.12.3. A full list of identified peptides is provided in the supplemental table. The four proteases used for protein digestion were thermolysin (THE) and chymotrypsin (CHY).

Peptide		Sequences	Proteases	Residues	HBS
FGF20	1	YLGMNDK(biotin)GEL	(CHY)	118-127	1
	2	YGSEK(biotin)LTSECIF	(CHY)	128-139	1
	3	K(biotin)HGDTGRRYF	(CHY)	157-166	1
	4	LNK(biotin)DGTPRDGARSK(biotin)RHQK(biotin)FTH	(THE)	169-189	1
	5	K(biotin)DLLMYT	(CHY)	205-211	3

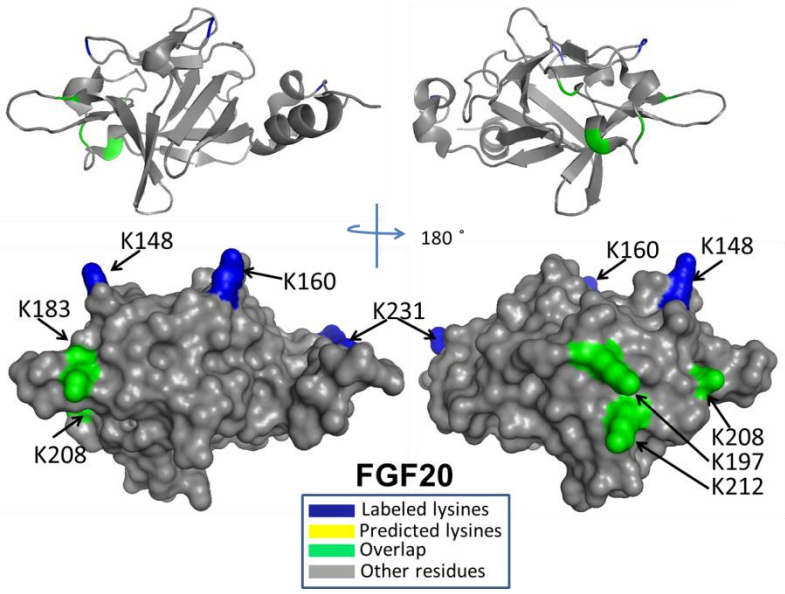


Figure 11 Position of biotinylated peptides in FGF-20 (residues 33-178) identified by structural proteomics mapped onto three dimensional structure（PDB 3F1R) [202]. Protect and label was used to identify the heparin binding sides of FGFs. Labelled peptides are coloured in *blue* and predicted lysines are coloured in *yellow*. The peptides overlapping with aligned HBS lysines are coloured in *green.*

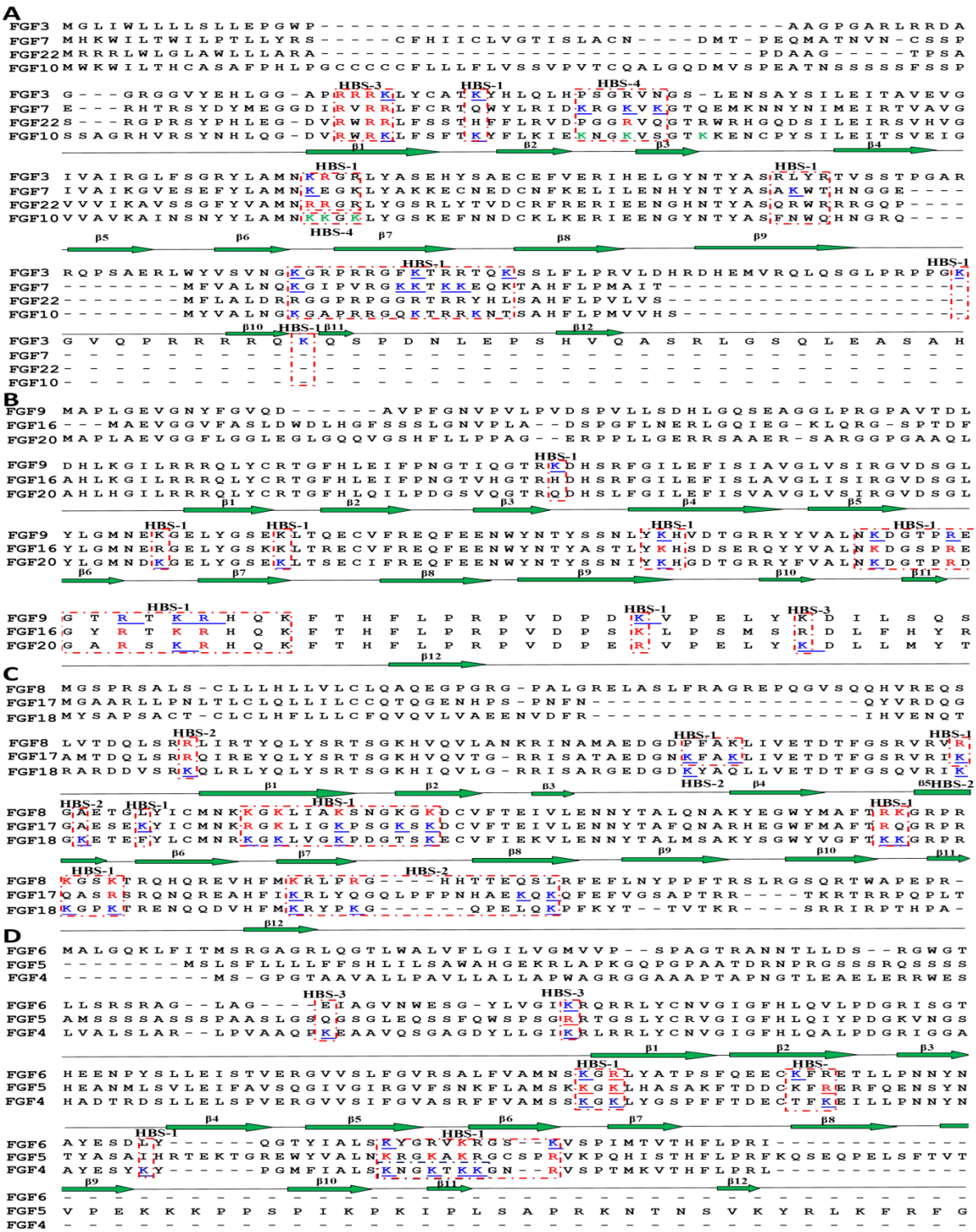


Figure 12 Sequence alignment of FGF subfamilies: The sequences were aligned with ClustalX and Dendroscope [203, 204]. Labelled lysines and overlapped with aligned HBS lysines are coloured in blue, predicted lysines and arginines are coloured in red. The aligned HBSs are boxed in red. Peptides of both acetylated and biotinylated lysines are coloured in green.

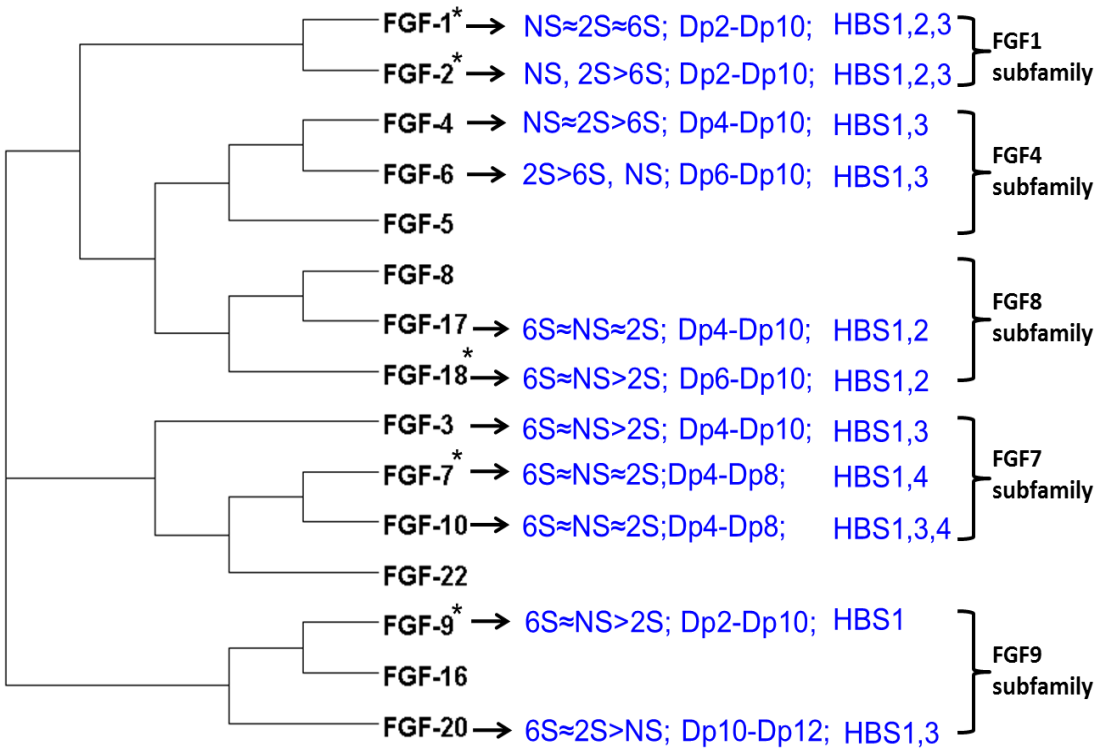


Figure 13 The heparin structural preference of FGFs and their heparin binding sites. *, data from Xu *et al* [34].

Supplemental data

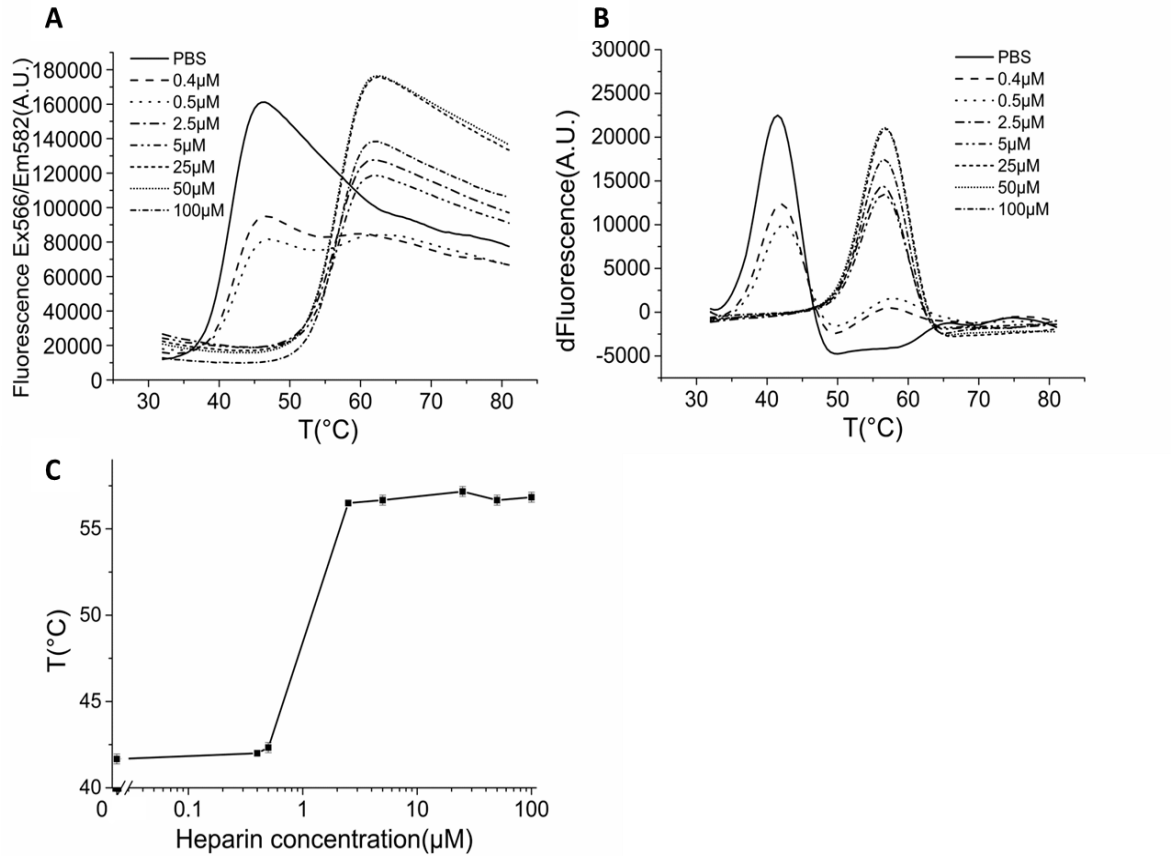


Figure S1 Stabilization effect of heparin on FGF10. Differential scanning fluorimetry of 5 μM FGF10 in the presence of various concentrations of heparin. A, melting curve profiles of FGF-10 (5 μM) with a range of heparin concentrations (0-100 μM). B, first derivative of the melting curves of FGF-10 in A. C, peak of the first derivative of the melting curves from B, which is the melting temperature, Tm (mean of triplicate ±S.E.).

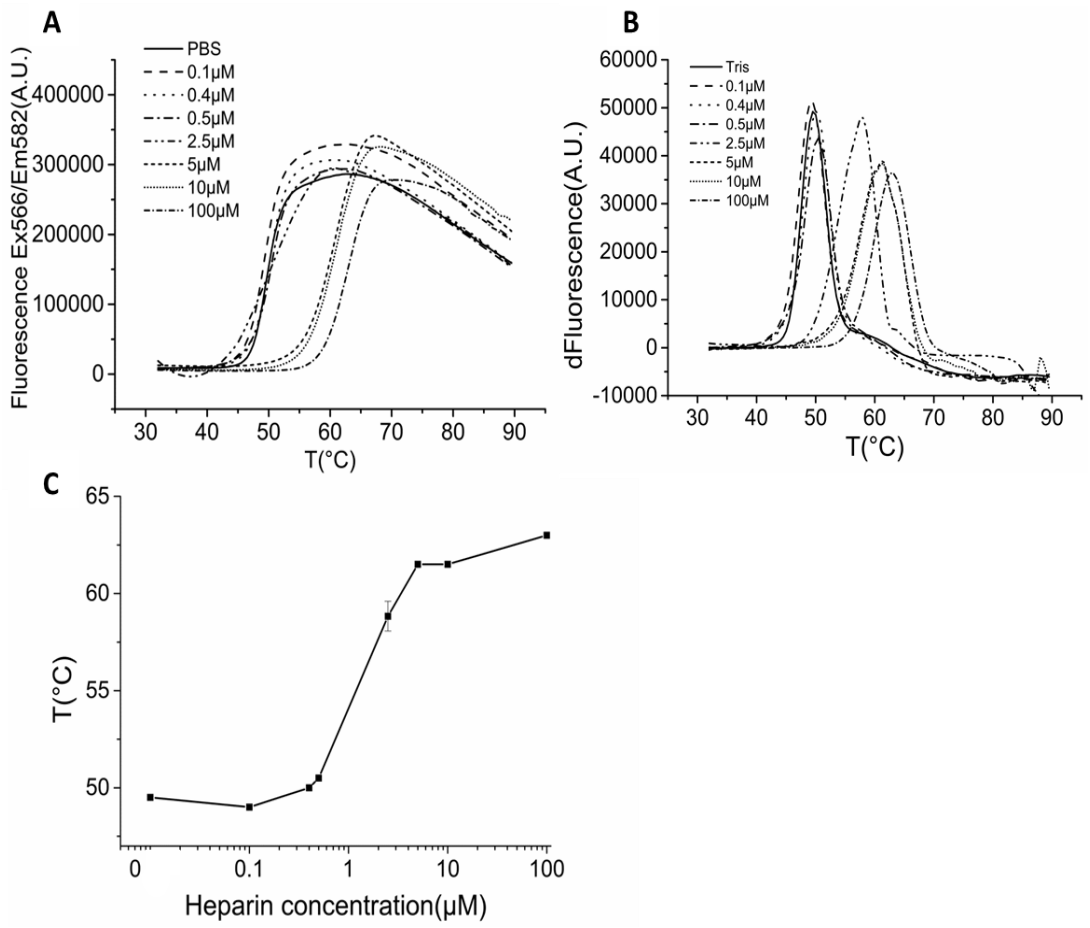


Figure S2 Stabilization effect of heparin on FGF4. Differential scanning fluorimetry of 5 μM FGF4 in the presence of various concentrations of heparin. A, melting curve profiles of FGF-4 (5 μM) with a range of heparin concentrations (0-100 μM). B, first derivative of the melting curves of FGF-4 in A. C, peak of the first derivative of the melting curves from B, which is the melting temperature, Tm (mean of triplicate ±S.E.).

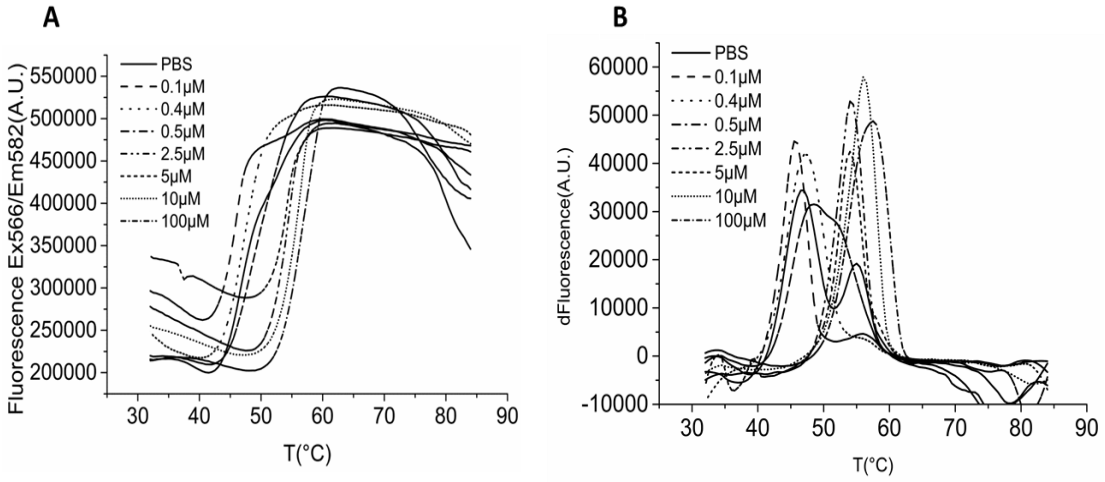


Figure S3 Stabilization effect of heparin on HT-FGF6. Differential scanning fluorimetry of 5 μM HT-FGF6 in the presence of various concentrations of heparin. A, melting curve profiles of HT-FGF6 (5 μM) with a range of heparin concentrations (0-100 μM). B, first derivative of the melting curves of HT-FGF6 in A.

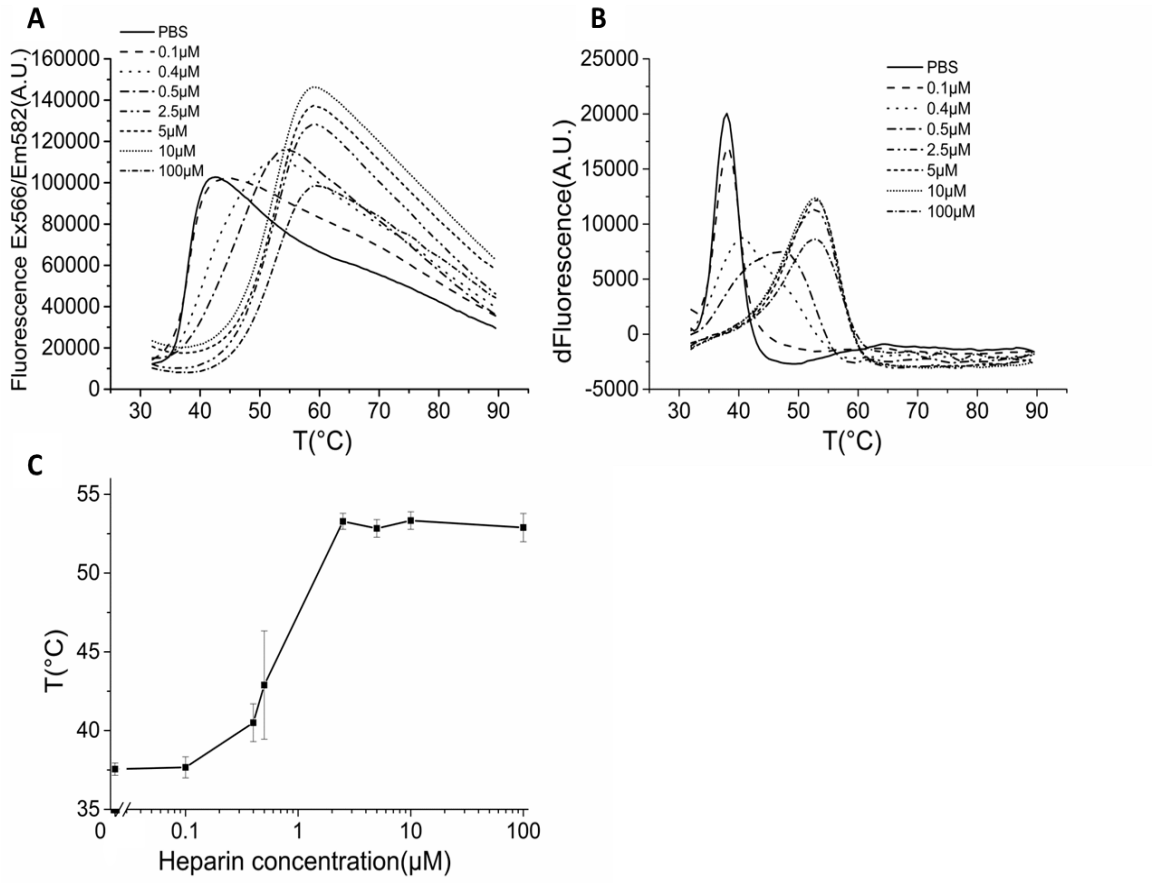


Figure S4 Stabilization effect of heparin on FGF17. Differential scanning fluorimetry of 5 μM FGF17 in the presence of various concentrations of heparin. A, melting curve profiles of FGF-17 (5 μM) with a range of heparin concentrations (0-100 μM). B, first derivative of the melting curves of FGF-17 in A. C, peak of the first derivative of the melting curves from B, which is the melting temperature, Tm (mean of triplicate ±S.E.).

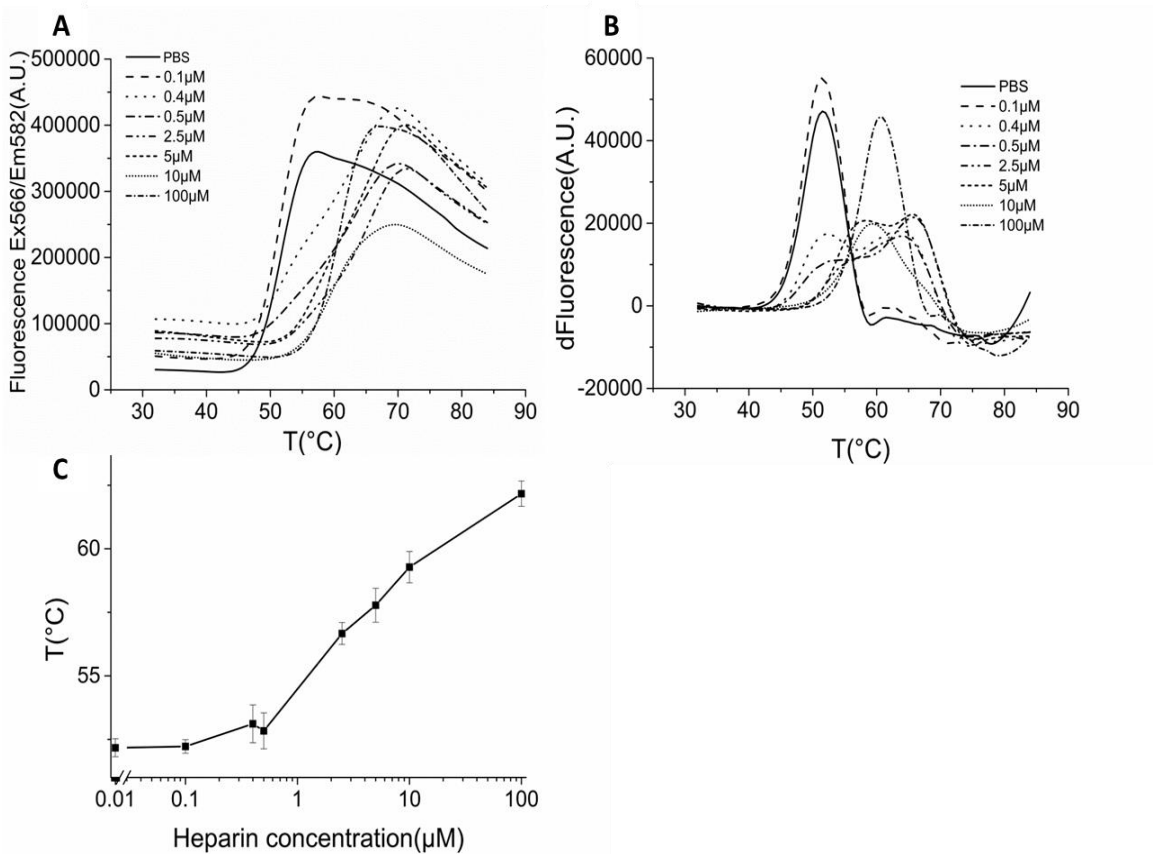


Figure S5 Stabilization effect of heparin on FGF20. Differential scanning fluorimetry of 5 μM FGF20 in the presence of various concentrations of heparin. A, melting curve profiles of FGF-20 (5 μM) with a range of heparin concentrations (0-100 μM). B, first derivative of the melting curves of FGF-20 in A. C, peak of the first derivative of the melting curves from B, which is the melting temperature, Tm (mean of triplicate ±S.E.).

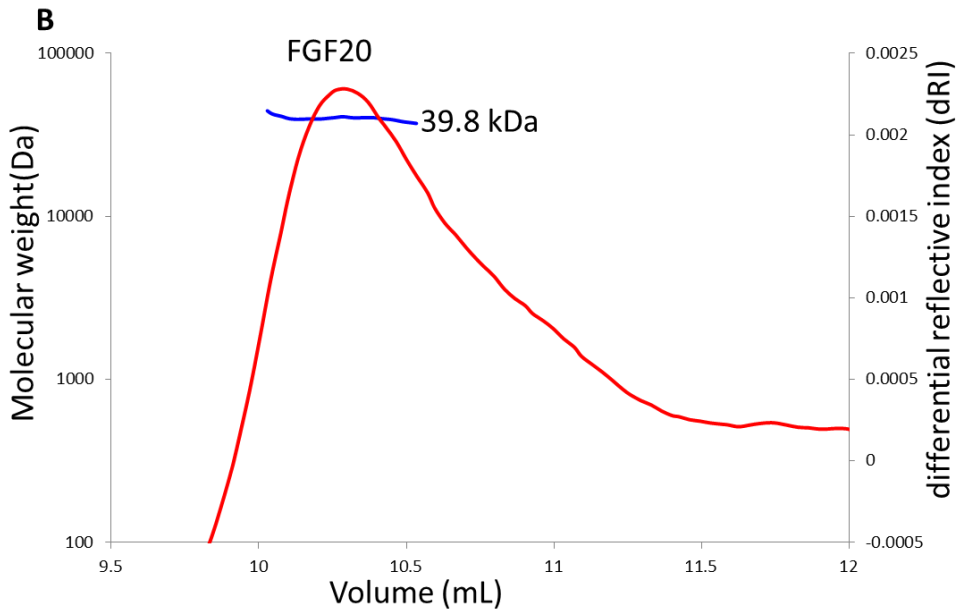


Figure S6 The average molecular mass per volume unit volume and the differential refractive index of FGF20. The red line shows the different reflective index (dRI). The blue line indicates the molecular mass of FGFs.

Table S1

Summary of peptides of FGFs identified by lysine targeted Protect and Label structure proteomics

Labelled peptides were identified by MALDI-Q-TOF and analysed by MS-digest of the package of ProteinProspector v 5.12.3. A full list of identified peptides is provided in the supplemental table. Four proteases used for protein digestion were trypsin (TRY), thermolysin (THE), chymotrypsin (CHY), Glu-c (GLU) and mixture of trypsin and Glu-C (TG).

Peptide		Sequences	Proteases	Residues	HBS
FGF3	1	YC(Carbamidomethyl)ATK(biotin)YHLQ	(THE)	49-57	1
	2	RTQK(biotin)SSLFLPR	(TRY)	171-181	1
FGF10	1	FSFTK(biotin)Y	(CHY)	83-88	1
FGF17	1	AK(biotin)LIVETDTF	(CHY)	84-93	1
	2	AHFIK(biotin)R	(TG)	172-177	2
	3	EAHFIK(biotin)R	(TG)	171-177	2
	4	LYQGQLPFPNHAEK(biotin)QK(biotin)QFE	(THE)	178-196	2
	5	K(biotin)QK(biotin)QFE	(GLU)	191-196	2
FGF4	1	VAMSSK(biotin)GK(biotin)	(CHY)	137-145	1
	2	IALSKNGKTKKGNRVSPT(3biotin/1acetyl)	(THE)	179-196	1
	3	AGDYLLGIK(biotin)R	(THE)	73-82	3
FGF6	1	QGTYIALSK(biotin)Y	(CHY)	177-186	1
	2	ATPSFQEEC(carbamidomethyl)K(biotin)F	(CHY)	149-156	3

Chapter 5 Characterization of HBS mutants of FGF2

5.1 Introduction

The heparin binding surfaces of FGF2 have been identified and classified into three heparin binding sites (HBS) [42, 170, 205]. The canonical heparin binding site (HBS1) is formed by the residues located on strand β1/ β2 loop, strand β10/ β11 loop, strand β11 and strand β11/ β12 loop [170]. The secondary HBS2 of FGF2 was identified and located to the region ^{115}YRSRKYSSWYVA^{126} [42, 205], whilst a third heparin binding site (HBS3) in FGF2 has been found at the N terminus of strand β1 [170]. HBS2 has been suggested on the basis of competition by peptides to bind heparin more strongly than HBS3 [42]. The peptide ^{117}SRKYSSWYVA^{126} corresponding to HBS2 was found to have low affinity for heparin (K_D, 120 ± 50 μM) and it was not able to inhibit the bioactivity of FGF2 [205]. The significant reduction in the mitogenic activity caused by truncation of the N terminus of FGF2 might show an important role of HBS3 [206]. However, this truncation removed part of strand β1 and likely destabilised the protein, which would be expected to affect the mitogenic activity.

To characterize the secondary HBSs and gain an insight into their function, two cDNAs were constructed coding for FGF2 mutants (HBS2 and HBS3). The lysines identified in HBS2 or HBS3 and adjacent arginines were replaced with alanine. This would remove the primary driver of the protein-polysaccharide interaction: ionic bonding of the side chains of these basic residues with sulfate and carboxylic acid on the polysaccharide. The effect, if any, of the mutations on the heparin structures required for binding was measured by differential scanning fluorimetry (DSF). Even the Tm stabilisation of FGF-2 is not directly related to activity, the Tm is used as

means of detecting binding, not as some predictor of activity [196]. Measurements of phosphorylation of p42/44^{MAPK} mitogen-activated protein kinase and cell growth were also performed to determine effects on the biological activity of the HBS mutations.

5.2 Methods

5.2.1 Production of FGF2 mutants

The cDNAs encoding FGF2 secondary HBS mutants were designed: FGF2 (HBS2) (208K/A, 219K/A, 249R/A, 251R/A, 252K/A) and FGF2 (HBS3) (160K/A, 163K/A, 164R/A). The detail of sub-cloning, expression and purification of FGFs from cDNAs is described in Chapter 3.

5.2.2 Differential Scanning Fluorimetry (DSF)

The melting temperatures of FGF2 mutants (HBS2 and HBS3) were determined with heparin (0-500 µM), heparin-derived oligosaccharides of different degrees of polymerization (dp), from dp 2 to dp 12, chemically modified heparin derivatives D1-D9, porcine mucosal HS, HA, CS, DS, and cation modified heparin as described in Section 2.9.

5.2.3 Phosphorylation of p42/44^{MAPK}

The stimulation of phosphorylation of p42/44^{MAPK} was measured in Rama (rat mammary) 27 fibroblasts, as described [171]. Briefly, cells were cultured until near confluent in 10 cm diameter dishes. After washing with sterile phosphate-buffered saline (PBS), the dishes were incubated with 10 mL step-down medium (DMEM

with 0.75 % (w/v) sodium bicarbonate, 20 mM L-glucosamine and 250 µg/mL BSA) for 24 h and then replaced with fresh step-down medium (9.5 mL) for another three hours. FGF2, FGF2 (HBS2) and FGF2 (HBS3) (final concentration: 200 pg/mL) were added to Rama 27 cells and incubated for 15 min and 1h. Afterwards, the cells were washed with ice-cold PBS twice and then collected by scraping in lysis buffer (20 % (v/v) glycerol, 4 % (w/v) SDS, 10 % (v/v) 2-mercaptoethanol in 120 mM Tris-Cl, pH 6.8 and coloured with 5 % (w/v) beta-bromophenol blue). The cell lysates were heated for 10 min at 98°C.

5.2.4 MTT assay

The Rama (rat mammary) 27 fibroblasts was incubated in 10 cm diameter dishes until near confluent. After detaching from dishes with trypsin, the cells were seeded in a 96-well plate at a density of 3,000 cells per well and cultured for 24 h at 37°C. Cell medium was replaced by 100 µL fresh SD medium and incubated for 24 h at 37°C. FGF2, FGF2 (HBS2) and FGF2 (HBS3) were then added to a final protein concentration of 10 ng/mL. In some cases heparin or dp 12 oligosaccharide were also added, as indicated in the figure legends. After 48 h incubation, 5 µL MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide) reagent (5 mg/mL) was added for 4 h. The dye was solubilized by the addition of 50 µL 10 % (w/v) SDS-Cl solution overnight at 37°C. The absorbance of each well was read in a spectrometer at 570 nm. Results were the mean values of experiments in quadruplicate.

5.3 Results and Discussion

5.3.1 The sugar structure required for binding to FGF2 mutants

The measurement of thermal denaturation by DSF of FGF2 and its HBS2 and HBS3 mutants in the presence of different concentrations of heparin concentrations (0- 100 μM) showed a concentration-dependent effect on the thermal stability of all three proteins. The melting temperatures of FGF2 (HBS2) and FGF2 (HBS3) were 35°C and 45°C, respectively, while that of wild-type FGF2 (around 55°C) was considerably higher (Fig. 5.2) and similar to that found previously [190]. Thus, both of the FGF2 mutants are less stable than wild-type FGF2 in the absence of heparin and the melting temperature of FGF2 (HBS2) is even lower than body temperature. One explanation is that the replacement of lysine and arginine with alanine (HBS2, K208A and R249A; HBS3, K163A and R164A), may create hydrophobic pockets, since the basic residues have longer alkyl chains than alanine. For example, a potential hydrophobic pocket is formed by the replacement of Arg-249 by alanine in the FGF2 (HBS2) (Fig. 5.1).

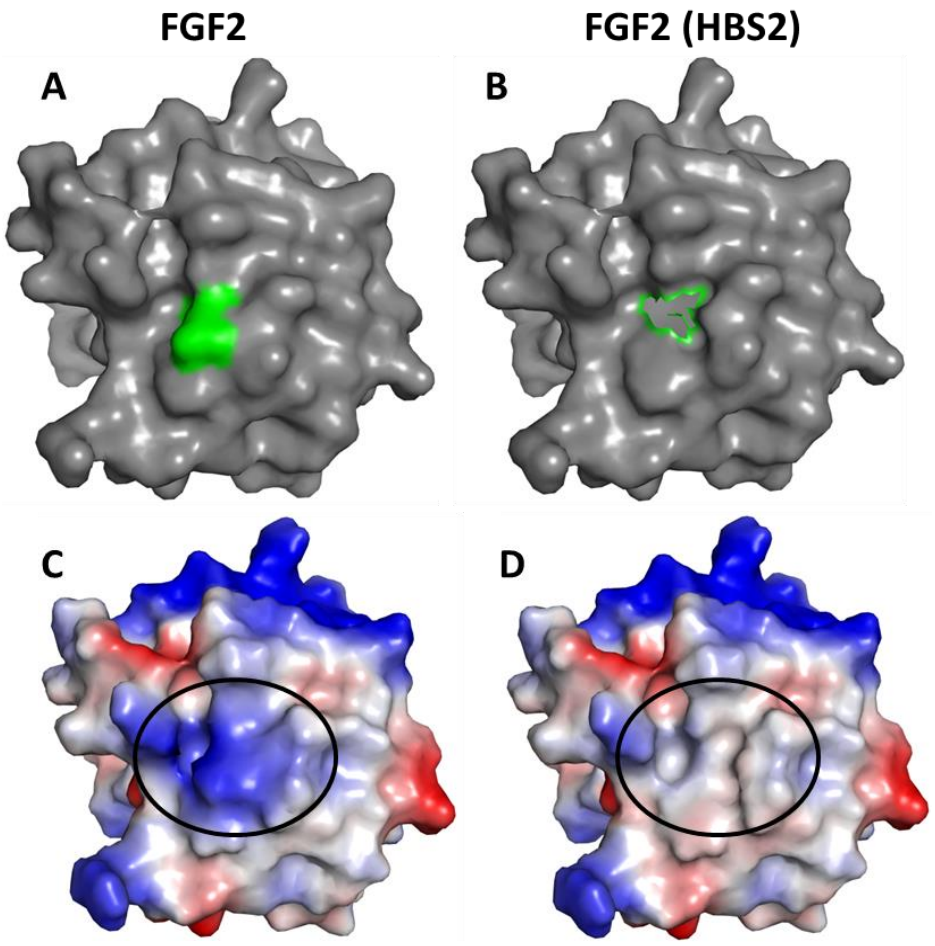


Figure 5.1 The crystal structure of FGF-2 (residues 143-288; PDB 2FGF) (Zhang et al., 1991). The green highlighted area in FGF2 and FGF2 (HBS2) are Arg-249 (A) and Ala-249 (B), respectively. C and D are the surface electronic distribution of FGF2 and FGF2 (HBS2), respectively. Figure rendered using Pymol, with the coordinates of 2FGF used to generate the structure of the FGF2 R249-A mutant.

In the presence of heparin, the thermal stability of wild-type FGF2 was increased to 78°C, which is similar to previous measurements [190]. The two HBS mutants were also stabilised by heparin, which is not surprising, since they still possess an intact HBS1. The melting curves show that heparin stabilised FGF2 (HBS2), starting from 0.4 μM heparin (heparin-FGF molar ratio, 12.5:1) and reaching the maximum melting temperature at 100 μM heparin (heparin-FGF molar ratio: 10:1) (Fig. 5.2). Although the stabilising effect of heparin on FGF2 (HBS3) was similar to the effect on FGF2 (HBS2), beginning at 0.4 μM heparin, it reached a maximum at 5 μM heparin (heparin-FGF molar ratio: 1:1) and did not change at higher concentrations of heparin. Since the maximum melting temperatures attained in the presence of heparin of FGF2, FGF2 (HBS2) and FGF2 (HBS3) were 78°C, 75°C and 80°C, heparin is able to stabilise FGF2 and its mutants to similar extents, though the lower initial stability of FGF2 (HBS2) seems to reduce the stability attained in the presence of heparin.

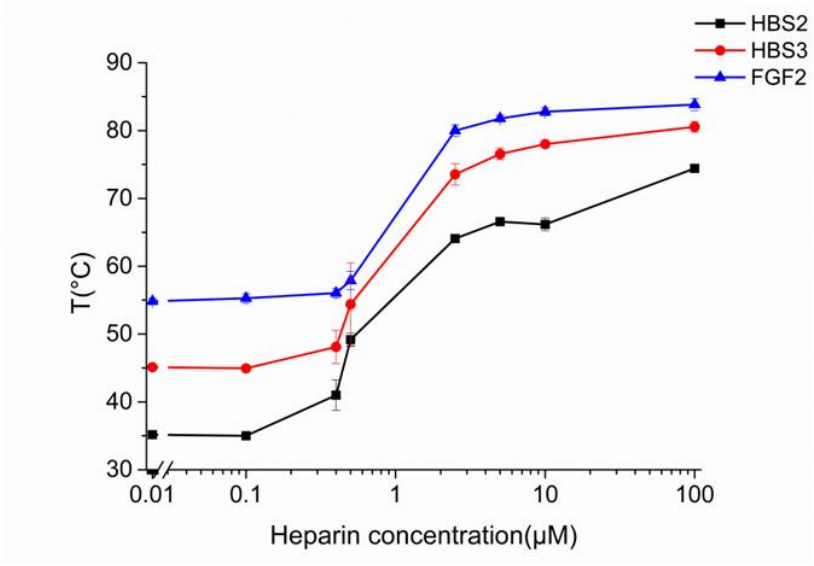


Figure 5.2 Melting temperature curves of FGF2 (blue), FGF2 (HBS2) (black) and FGF2 (HBS3) (red). Differential scanning fluorimetry of 5 μM FGF2 (blue), FGF2 (HBS2) (black) and FGF2 (HBS3) (red) in the presence of various concentrations of heparin (0-100 μM). Results are the mean of triplicates ±S.E.

The effect of mutating HBS2 and HBS3 on the specificity of FGF2 for different sugar structures was then explored. In the case of FGF2 (HBS2), both the single desulfated heparin D2 (2-O-sulfate and 6-O-sulfate) and D4 (2-O-sulfate and N-sulfate) stabilised FGF2 (HBS3) to 80 % and 100 %, respectively, of the level observed with heparin respectively, whereas D3 (6-O-sulfate and N-sulfate) only stabilised FGF2 (HBS3) half as effectively as heparin (Fig. 5.3A). Although the thermal stabilisation effect of any doubly desulfated heparins with either 2-O-sulfate or 6-O-sulfate on FGF2 (HBS2) was less than that seen with the singly desulfated heparins, the heparin with just N-sulfate (D7) achieved a stabilisation effect on FGF2 (HBS2) similar to HS and D3. There was a slight binding of FGF2 (HBS2) to DS, but no detectable interaction to HA or to CS. FGF2 (HBS2) did not interact with dp 4 and a dp 6 was the minimum size of oligosaccharide required for stabilising FGF2

(HBS2). A dp 12 oligosaccharide imparted the same stabilisation to FGF2 (HBS2) as full length heparin (Fig. 5.3B). FGF2 (HBS2) did not distinguish between the different cationic forms of heparin (Fig. 5.3C). Thus, both FGF2 (HSB2) and wild-type FGF2 prefer to bind to the combination of 2-O-sulfate and N-sulfate than to other singly desulfated structures, however, they have different preferences for singly sulfated structures. FGF2 has a weak interaction with N-sulfated heparin and no interaction with 6-O-sulfated heparin, whereas FGF2 (HBS2) has a considerably stronger interaction with N-sulfated heparin and a weak interaction with 6-O-sulfated heparin.

FGF2 (HBS3) was more effectively stabilised by any of the singly desulfated heparins (D2, D3 and D4) than any of the doubly desulfated heparins (D5, D6 and D7) (Fig. 5.3A). The single desulfated heparin D2 and D3 stabilised FGF2 (HBS3) to 57 % and 34% of the level observed with heparin respectively, whereas D4 (2-O-sulfate and N-sulfate) stabilised FGF2 (HBS3) to the same extent as heparin. Therefore, FGF2 (HBS3) prefers to bind structures containing 2-O-sulfate and N-sulfate, such as wild-type FGF2 [40, 160, 190]. For singly sulfated heparins (D5, D6 and D7), there was only a small stabilising effect on FGF2 (HBS3). Similar to D2 and D3, HS stabilised FGF2 (HBS3) to 64 % of the level seen with heparin. However, FGF2 (HBS3) did not show a detectable interaction with HA or CS, but DS clearly does bind, albeit not as well as HS. The minimum size of oligosaccharide required to stabilise FGF2 (HBS3) was dp 8. A dp 12 could stabilise FGF2 (HBS3) as effectively as full length heparin (Fig. 5.3B). FGF2 (HBS3) did not bind differently to any of the different cationic forms of heparin like FGF2 (HBS2) (Fig. 5.3C). Thus, while FGF2 (HBS3) binds similar structures as wild-type FGF2, it required a

substantially longer oligosaccharide (wild-type FGF2 binds dp 4, whereas FGF2 (HBS3) required a dp 8 as the minimum binding structure).

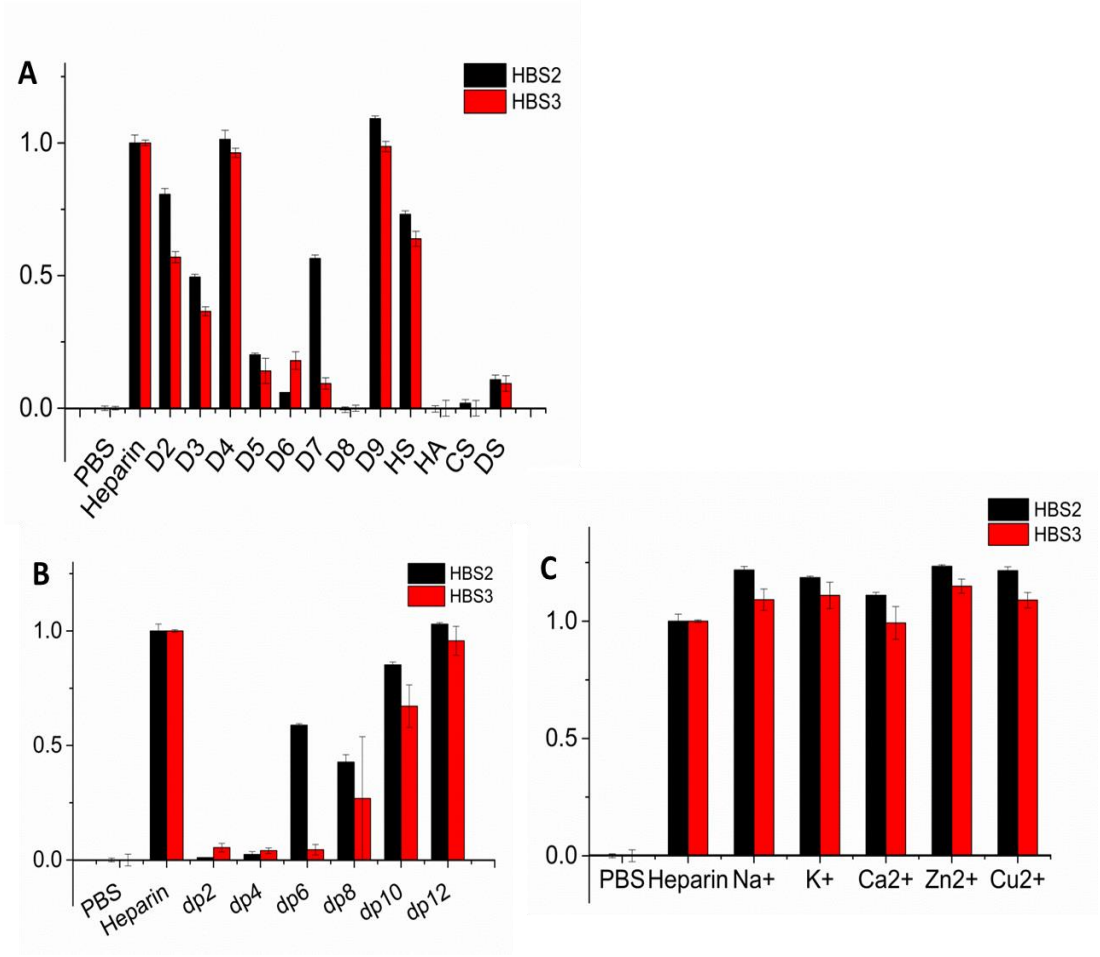


Figure 5.3 Differential scanning fluorimetry analysis of binding of GAG derivatives to FGF2 (HBS2) (black) and FGF2 (HBS3) (red). Differential scanning fluorimetry was performed with a range of heparin-based poly- and oligosaccharides with 5 μM protein and 175 μg/ml sugar and the thermal stabilisation relative to the PBS control (=0) and heparin (=1) was calculated. A, thermal stabilisation effect of chemically modified heparins (D2–D9), and other glycosaminoglycans (HS, HA, CS, and DS); B, heparin-derived oligosaccharides, ranging from dp 2 to dp 12; C, cation-modified heparin forms. Results are the mean of triplicates after normalization ±S.E., and an apparent absence of an error bar is due to a small S.E.

5.3.2 Biological activities of FGF2 and its HBS mutants on Rama 27 fibroblasts

The biological activities of the HBS2 and HBS3 mutants of FGF2 were measured and compared to those of wild-type FGF2. Since FGF2 prefers to bind FGFR 1c and FGFR 3c [35], FGF2 and its mutants were tested on Rama 27 fibroblasts which express FGFR 1c [207]. After stimulation of quiescent Rama 27 cells with FGF2 (200 pg/mL) for 10 min or 60 min, two strong bands corresponding to dually phosphorylated p42/44^{MAPK} were observed after 10 min, which had decreased in intensity by 60 min, as seen previously [121, 123, 208] (Fig. 5.4A). The mutants (HBS2, HBS3) also resulted in a similar stimulation of phosphorylation p42/44^{MAPK} in cells after 10 min (Fig. 5.4A). However, after 60 min stimulation with FGF2 (HBS2), the level of phosphorylation of p42/44^{MAPK} had returned to basal levels (Fig. 5.4A). Thus, FGF2 (HBS2) seemed not to promote a sustained plateau of phosphorylation p42/44^{MAPK} after the initial early peak, in contrast to wild-type FGF2 (Fig. 5.3A [121, 123, 208]) and to FGF2 (HBS3). This may reflect the lower stability of FGF2 (HBS2), shown by DSF (Fig. 5.2). Since heparin stabilises FGF2 (wild type and mutants), the same ratio of heparin: FGF2, which was found to maximally stabilise FGF2, was applied to cells to determine if the low stability of FGF2 (HBS2) was the reason for its inability to stimulate the phosphorylation p42/44^{MAPK} for 60 min. The activity of FGF2 was equivalent to that of FGF2 with heparin, which showed that heparin did not affect the stimulation of phosphorylated p42/44^{MAPK} (Fig. 5.4B) [121, 178]. In the presence of heparin, which stabilises FGF2 (HBS2) to the same extent as heparin (Fig. 5.4B), the stimulation of phosphorylation p42/44^{MAPK} at 60 min by FGF2 (HBS2) was now apparent. To explore in more depth the effect of stability of FGF2 (HBS2) on the stimulation of phosphorylation of p42/44^{MAPK} on Rama 27 cells, a more complete kinetic analysis was performed (Fig. 5.4C). The phosphorylation of

p42/44^{MAPK} reached a maximum at 5 min after the addition of FGF2 and the initial peak of phosphorylation declined to a sustained plateau between 30 min and 60 min. When Rama 27 cells were stimulated with FGF2 (HBS2), the level of phosphorylation of p42/44^{MAPK} also reached maximum levels after 5 min, but then declined from 20 min to 60 min, at which time it was still above basal level. Thus, there were still differences in the sustained phosphorylation of p42/44^{MAPK} by FGF2 (HBS2) and wild-type FGF2. These may be due to the mutation of HBS2. However, it is also possible that the amount of heparin used to stabilise the FGF2 (HBS2) was insufficient, because there are considerable numbers of heparin binding sites on cells [209], which would therefore decrease the effective concentration of free heparin available to stabilise the mutant FGF2.

The stimulation on cell growth of Rama 27 fibroblast by wild-type FGF2 and FGF2 (HBS3) was measured using the MTT assay. In contrast to non-stimulated cells, the wild-type FGF2 and FGF2 (HBS3) stimulated the incorporation of MTT by almost two fold (Fig. 5.4). Even though FGF2 (HBS3) is less stable than wild type FGF2, it is clearly sufficiently stable to stimulate cell growth to almost the same extent as wild type FGF2. This indicates that HBS3 does not directly affect the intracellular signalling process that is required for stimulation of cell growth by FGF2.

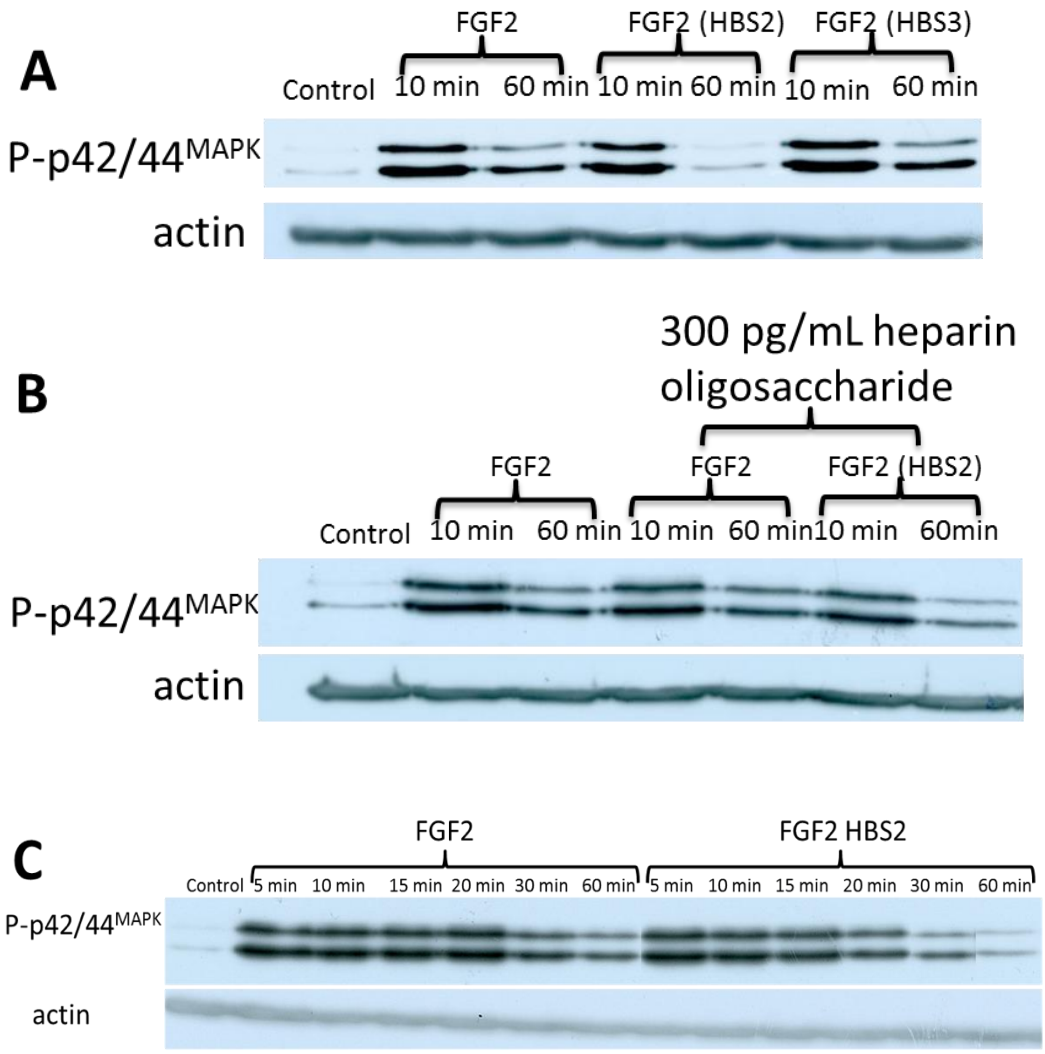


Figure 5.4 Kinetics of phosphorylation of p42/44 ^{MAPK} induced by FGF2 and its mutants. Cells were grown in 10 cm diameter dishes and incubated in step-down medium for 24 h as described in Section 5.2.2. *A*, stimulation of p42/44^{MAPK} phosphorylation by FGF2, FGF2 (HBS2) and FGF2 (HBS3). FGF2 and its mutants (200 pg/mL) were applied to cells for 10 min or 60 min.; *B*, stimulation of p42/44^{MAPK} phosphorylation by FGF2 and FGF2 (HBS2) (200 pg/mL) with 300 pg/ mL heparin and FGF2 without heparin. *C*, Time-course of stimulation of p42/44^{MAPK} phosphorylation by FGF2 and FGF2 (HBS2) between 0 min and 60 min. The blots were re-probed with anti-actin to indicate the level of gel loading.

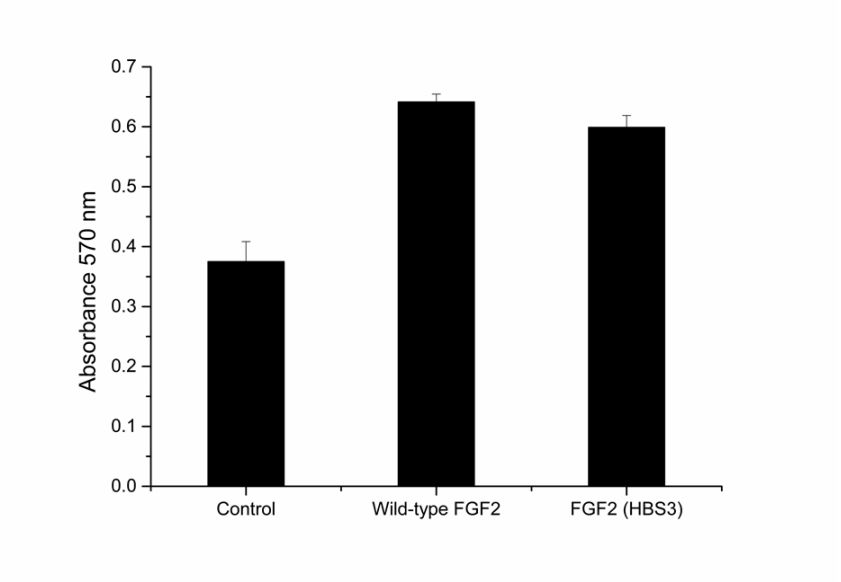


Figure 5.5 Incorporation of MTT by Rama 27 fibroblasts stimulated by FGF2 or its mutant. Cells were seeded in 96 wells plate at a density of 3,000 cells per well, as described in the methods. After incubation in SD medium for 24 h, the cells were treated with 10 ng/mL wild-type FGF2 and FGF2 (HBS3) for 48 h. The cells were mixed with 5 µL MTT reagent (5 mg/mL) for 4 h and then stop solution (50 µL 10 % (w/v) SDS-Cl) was added overnight. The absorbance was read at 570 nm in a spectrometer. The control groups are cells treated with SD medium only. Results are the mean ± S.E of quadruplicate wells.

5.4 Conclusions

Two FGF2 mutants (HBS2 and HBS3) were constructed and produced to gain an insight into the secondary HBSs and their function. The secondary binding site HBS2 and HBS3 were mutated by the replacement of lysines and arginines with alanines in FGF2 (HBS2) and FGF2 (HBS3), respectively. Both mutants were charaterised by differential scanning fluorimetry (DSF) to determine the various features of heparin structures required for optimum binding; as measured by the effect of oligosaccharide binding on the Tm of the FGFs. Both of the FGF2 mutants were found to have a lower melting temperature than wild-type FGF2 in the absence of heparin. The lower stability particularly of FGF2 (HBS2) may be due to the choice of alanine as a substitute for lysines and arginines, which may result in the formation of hydrophobic pockets (Fig. 5.1). When compared to wild-type FGF2, the classic sulfation pattern required for binding to heparin [40, 41, 160] of the two mutants was not changed, which indicated the driving role of HBS1 in the interaction between FGF2 and heparin/HS. However, binding of the mutants to heparin structures with lower levels of sulfation or of restricted length was not identical to that seen with wild-type FGF2. For example, FGF2 (HBS2) bound more strongly to N-sulfated heparin compared to wild-type FGF2. Both FGF2 (HBS2) and FGF2 (HBS3) required a larger oligosaccharide for minimal binding than did wild-type FGF2. It is not clear to what extent these differences are due to an absence of HBS2 and HBS3 between the mutants and the wild type protein. One interpretation of these differences is that they reflect a lower rebinding capacity for the polysaccharide by the mutants. Thus, in the case of oligosaccharides, following dissociation, they might rapidly rebind the FGF2. There is some evidence for this from protect and label

experiments, where some lysines are found to be both acetylated and biotinylated [170]. This is due to transient dissociation from the heparin of one lysine in the HBS, during which time it reacts with the protection reagent, NHS acetate. In the mutants, rebinding after more substantial dissociation (of more than one lysine) may involve a different HBS, so reducing the number of HBSs reduces the rebinding capacity. In the case of small oligosaccharides such as dp 6, their rebinding may involve their interaction with HBS2 or HBS3 after dissociation and then they could then migrate across the surface of the FGF2 back to HBS1. Such movement of the polysaccharide on a protein surface has been proposed for neuropilin-I [194]. The biological activities of FGF2 and FGF2 mutants were also measured in terms of their stimulation on the phosphorylation of p42/44^{MAPK} mitogen-activated protein kinase and cell growth, the latter reported by incorporation of MTT. FGF2 mutants could stimulate the phosphorylation of p42/44^{MAPK} in a similar manner as wild-type FGF2. However, the lower stability of FGF2 (HBS2) confounded this result, since alone it was unable to stimulate longer-term phosphorylation of p42/44^{MAPK}. By including heparin to stabilise FGF2 (HBS2), it was shown that the lack of stability of this mutant is likely to be responsible for this difference. Taken together with the similar results for p42/44^{MAPK} phosphorylation and cell proliferation induced by FGF2 and FGF2 (HBS3), the results suggest that HBS2 and HBS3 may not be involved directly in generating these cellular responses. In another study, the secondary binding sites of FGFs were shown to have the capacity to cross-link HS chains. This may affect the movement of the FGF2 in extracellular matrix [198]. In turn cross-linking of HS chains by FGF2 could result in the confined motion rather than diffusive motion, as observed in pericellular matrix [178]. This may have two consequences. First, the HBS2 and HBS3 mutants of FGF2 may diffuse differently in the pericellular matrix,

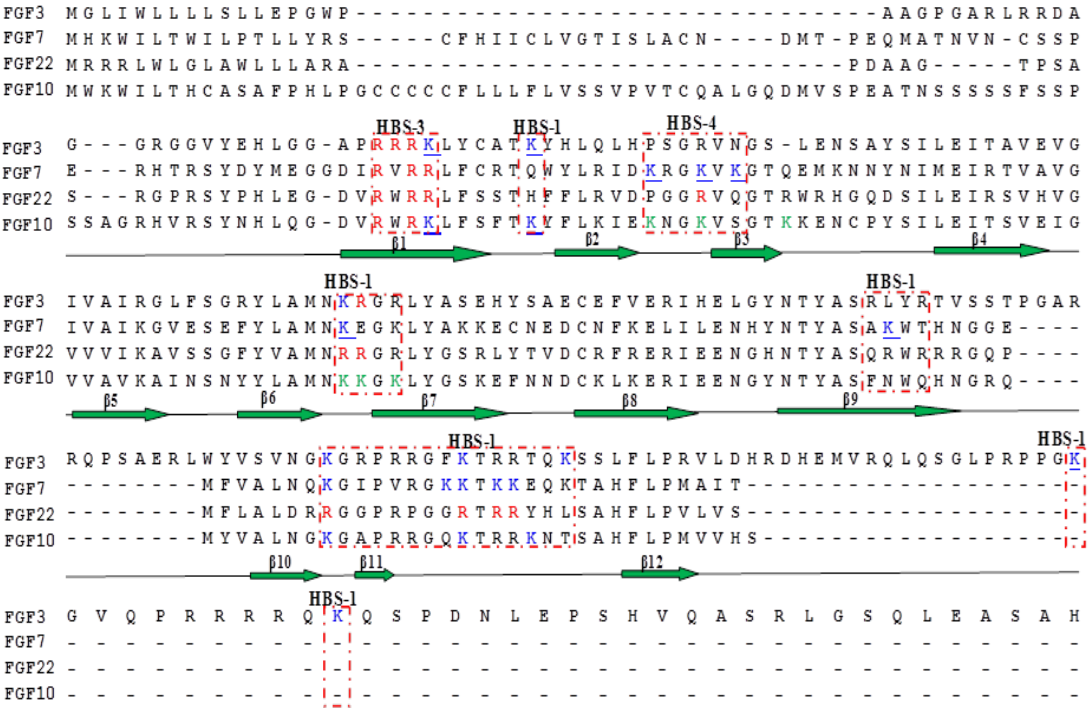
which would affect the types of gradients they are capable of forming. Secondary, because they may as a result be less susceptible to trapping in the extracellular matrix, in cell culture they would spend more time proportionally in the bulk culture medium compared to wild-type FGF2, where their lower stability (especially that of FGF2 (HBS2)) would result in their denaturation and loss of activity.

Chapter 6 Arginine-targeted protect and label

6.1 Introduction

The technique called "protect and label" used in Chapter 6 was developed by Ori *et al*., [170] to selectively label lysine side chains directly involved in heparin binding, that are protected by interaction with the polysaccharide. However, it is not robust for heparin binding sites which are rich in arginine, but have few or any lysines. For example, the predicted HBS1 of FGF22 includes one lysine, with the remainder of the basic side chains being contributed by arginine residues (Fig. 6.1). Indeed, in the FGF7 subfamily, the alignment of sequences considered to contribute to heparin binding sites indicates that there are many arginine residues. For the characterisation of such heparin binding sites, an arginine targeted technique is required, which is the focus of this Chapter. For completing the process of protection and labelling, two chemicals specific to the guanidino chain of arginine are needed. Phenylglyoxal (PGO) has been used as an arginine specific to reagent in the inhibition of chloride and sulfate equilibrium exchange across human red cells [210, 211]. A reversible bond is created when the glyoxal reacts with arginine residues at 2:1 ratio [212], whereas, an irreversible bond was formed when the ratio of glyoxal and arginine was 1:1 [213] (Fig. 6.2). Thus, there are potentially at least two products from this reaction. For the labelling part, another arginine specific reagent was used, 4-azidophenylglyoxal (APG), which has the same dicarbonyl moiety as phenylglyoxal. The inhibition by APG of enzymes which have arginine at active sites has demonstrated its ability to react with the arginine side chain [214], while the azido group of APG could be photolysed with UV light to produce an unstable nitrene that reacts with

nucleophiles, which was used in a cross-linking step [215]. The azido group of APG has also been shown to react with alkynes by click chemistry to form triazoles [216] (Fig. 6.3). Biotin DIBO, a strained cyclooctyne from Life Technologies would then be used as the alkyne for a copper-free click reaction with the azide structure and the biotin group would thus enable the purification of labelled peptide. Therefore, phenylglyoxal and triazole production were chosen for the development of an arginine targeted protect and label method.



Figures 6.1 Sequence alignment of the FGF7 subfamily: Lysines and overlapped with aligned HBS lysines are coloured in blue, arginines predicted from sequence alignment are coloured in red. The aligned HBSs are boxed in red. Peptides of both acetylated and biotinylated lysines are coloured in green (data from Chapter 4 and [34]). The sequences were aligned with ClustalX and Dendroscope.

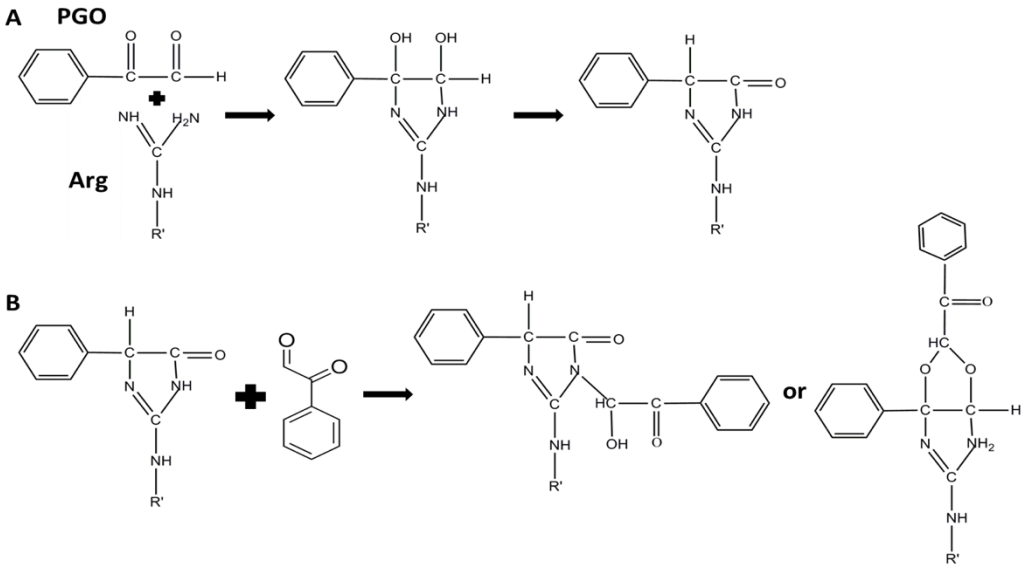


Figure 6.2 Reaction between arginine (Arg) and phenylglyoxal (PGO). A. the guanidyl group of arginine reacts with the dicarbonyl moiety of phenylglyoxal to form a Schiff's base; B. The complex can react with a further phenylglyoxal to form alternative structures at a ratio of Arg and PGO of 1:2 [214].

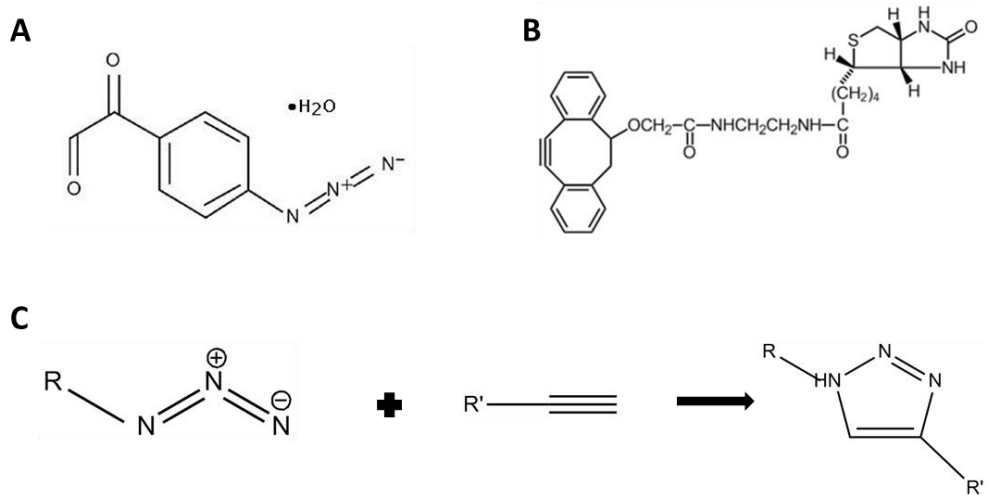


Figure 6.3 Reaction between *P*-azidophenylglyoxal (APG) and biotin DIBO alkyne. A. p-Azidophenylglyoxal monohydrate (APG); B. Biotin DIBO alkyne; C. The azide group of APG reacts with the alkyne of biotin DIBO to form a triazole structure by click chemistry.

6.2 Materials and methods

6.2.1 Materials

Peptides (FA, BK, NA) (Table 6.1) and phenylglyoxyl (PGO) were purchased from Sigma-Aldrich. Azidophenylglyoxal monohydrate (APG) was purchased from Bioworld (Bioquote Ltd, York, UK).

Table 6.1 The peptides used to develop arginine selective labelling. pGlu, pyroglutamyl peptide.

Peptide	Sequence	Molecular weight (kDa)
FA (Fibronectin Adhesion-promoting Peptide) :	WQPPRARI	1023
BF (Bradykinin Fragment 2-9) :	PPGFSPFR	904
NA (Neurotensin) :	pGlu- LYENKPRRPYIL	1672

6.2.2 Arginine targeted 'protect and label'

Phenylglyoxal monohydrate (PGO) was dissolved in DMSO (dimethyl sulphoxide) to 4 M as stock, and then adjusted to the required concentration with PBS (137 mM NaCl, 2.7 mM, KCl, 10 mM Na_2HPO_4 • 2 H_2O, 2.0 mM KH_2PO_4, pH 7.4). A series of concentrations of phenylglyoxal were reacted with peptide FA (1.5 μg), which includes two arginine residues for different reaction times at room temperature in the dark. Various concentration of p-azidophenylglyoxal (APG) were prepared in PBS and similarly reacted with peptide FA (1.5 μg). The labelled peptides were purified with C18 Zip Tips and analyses were performed on a MALDI-TOF mass spectrometer (Waters-Micromass, Manchester, UK) (Section 2.10 .5). Peptides were

analysed using a saturated solution of alpha-cyano-4 hydroxycinnamic acid in 50 % acetonitrile/0.1 % trifluoroacetic. Peptides were selected in the mass range of 1000-3500 Da. Data analysis was performed using Mascot.

6.3 Results

6.3.1 Method development for arginine protection by phenylglyoxal

The first step in arginine targeted 'protect and label' is the protection of arginines that are exposed to solvent, because, they are not involved in heparin binding. Since phenylglyoxal has been shown to specifically react with arginine side chains, it was chosen to block exposed arginines. First of all, a series of concentrations of phenylglyoxal were tested with peptide FA, which contains two arginines, for different reaction times in order to identify the conditions required for the reaction to go to completion (Table 6.2). Two reaction times (5 min and 30 min) were used, each with five different concentrations of phenylglyoxal. Peptide FA was unmodified after reaction with either 50 mM or 100 mM phenylglyoxal for 5 min (Figs 6.8A and B), which indicated that neither of these two concentrations was able to provide full protection over this time. Similarly, with 10 mM and 50 mM phenylglyoxal and 30 min reaction time (Figs 6.9A and B), not all of arginine residues reacted. However, phenylglyoxal at 200 mM modified all of the arginines in the FA peptide over 5 min and 30 min (Figs 6.8C and 6.9C). No peak corresponding to the peptide was identified when it was reacted with 1M phenylglyoxal (Fig. 6.9D), which could be due to the aggregation of peptide induced by the high concentration of phenylglyoxal.

Table 6.2 Summary of conditions used to react peptide FA with phenylglyoxal.

Phenylglyoxal (mM)	5 min	30 min
10	---	uFA, FA+PGO, FA+2xPGO
50	uFA, FA+PGO, FA+2xPGO	uFA, FA+PGO, FA+2xPGO
100	uFA, FA+PGO, FA+2xPGO	---
200	FA+PGO, FA+2xPGO	FA+PGO, FA+2xPGO
1000	---	None

Peptide FA was reacted with different concentrations of phenylglyoxal, 10 mM (Fig. 6.9A), 50 mM (5 min, Fig. 6.8A; 30 min, Fig. 6.9B), 100 mM (Fig. 6.8B), 200 mM (5 min, Fig. 6.8C; 30 min, Fig. 6.9C) and 1 M (Fig. 6.9D) at room temperature in the dark. FA, Fibronectin adhesion-promoting peptide; uFA, FA alone; FA+PGO, peptide FA was reacted at a molar ratio of 1:1 with PGO; PGO+2xPGO; peptide FA was reacted at a molar ratio of 1:2 with PGO. None, no peptide was detected. No peptide was detected when peptide FA was reacted with 1.0 M PGO for 30 min.

Phenylglyoxal at 200 mM could, over 5 min, react fully with the arginines in FA. Interestingly, two reaction products of peptide FA were detected: type A was FA modified by one PGO, where the ratio of arginine: PGO is 2:1; type B was FA modified by two PGO, which was the expected product at a ratio of arginine: PGO (1:1) (Fig. 6.4B). One interpretation is that PGO could not protect all the arginines when the product was of type A. However, the two arginine residues are close to each other and separated by just an alanine residue, which has a short side chain. Thus, another possibility could be that the dicarbonyl group of the PGO has reacted with the side chains of the two adjacent arginines to form a Schiff's base and hence

the arginines are fully protected. To distinguish these possibilities, modified FA was treated with trypsin. Unmodified peptide FA was successfully digested by trypsin and the large fragment 'WQPPR' was identified (Fig 6.4C). After trypsin digestion of modified FA, only the two peptides, which was detected in the absence of the digestion, were detected (Fig 6.4D). Thus, the peptides modified with PGO in the ratio of arginine: PGO 2:1 and 1:1 were protected from digestion, which means both of the arginines are subject to modification by PGO.

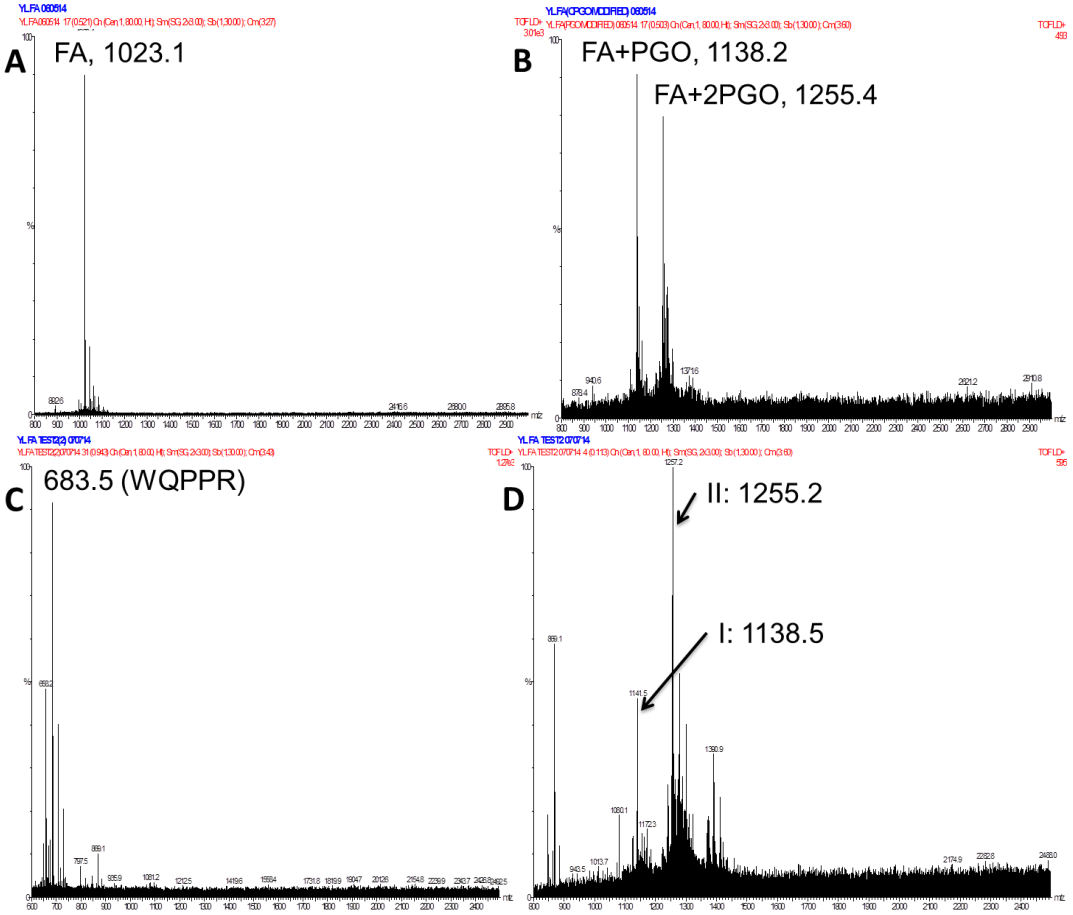


Figure 6.4 Mass spectra of the digestion of unmodified and modified FA. A. FA, Fibronectin Adhesion-promoting peptide; B. Peptide FA was reacted with 200 mM phenylglyoxal for 30 min. FA+PGO, peptide FA modified with one PGO; FA+2xPGO; peptide FA modified with two PGO. C. unmodified peptide FA digested with trypsin. The large part of the sequence was identified as 'WQPPR'. D. modified peptide FA digested with trypsin and only two modified peptides (I and II) were detected.

It is clear from the above that the position of arginine resides affects the reaction product. Peptide BF (Bradykinin Fragment), which has one arginine at its C terminus, was used to identify the effect of this position, where the arginine side chain is relatively unencumbered (Fig. 6.5A). This peptide was reacted with 200 mM phenylglyoxal in the dark for 30 min. Two types of modified peptide BF were identified: one peptide with the arginine modified by a single PGO (arginine: PGO, 1:1); a second peptide with the arginine modified by two PGO (arginine: PGO, 1:2) (Fig. 6.5B). Since the arginine at the C terminal has more free space for its side chain, the PGO could create a reversible bond at a 2:1 ratio of its dicarbonyl group to arginine as described in Takahashi's study [212]. To establish that this was the case, two peptides: peptide I (RPYIL, molecular mass 663.3) and peptide II (LYENKPR, molecular mass 1062.0) were produced by treating peptide NA with trypsin. This yields one peptide with arginine at the N terminus (peptide I) and another one with arginine at the C terminus (peptide II) (Fig. 6.6A). Both peptides were added to 200 mM PGO in the dark for 30 min. In each case they were modified with one PGO or two PGO (Fig. 6.6B).

The use of peptides with arginine residues in defined contexts demonstrated that phenylglyoxal was likely to fully protect all of the arginine residues in a protein regardless of the adjacent sequence. The position of arginine in peptides could affect the formation of the Schiff's base and multiple products are clearly produced. However, these are resolvable by mass spectrometry. The next step was to determine whether azidophenylglyoxal (APG) could similarly fully react with arginine residues.

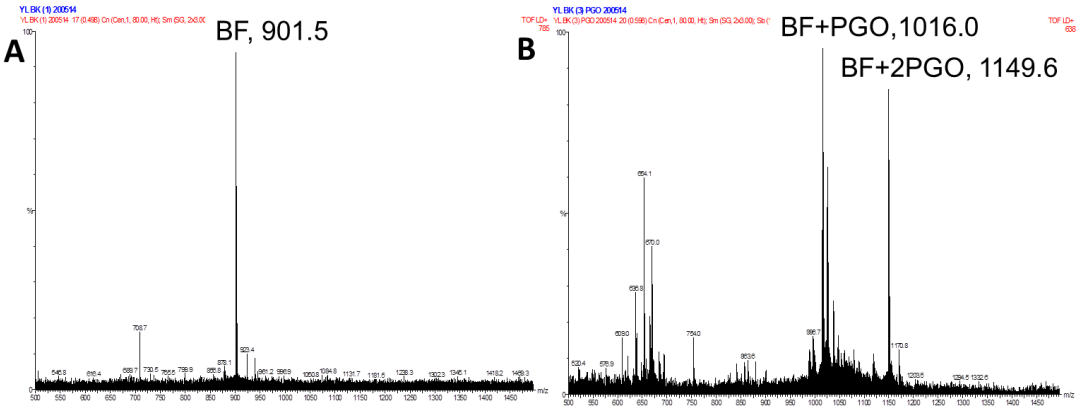


Figure 6.5 Mass spectra of BF reacted with 200 mM phenylglyoxal for 30 min.

A. BF, Bradykinin Fragment 2-9; B. BF+PGO, peptide BF modified with one PGO; BF+2xPGO; peptide BF modified with two PGO.

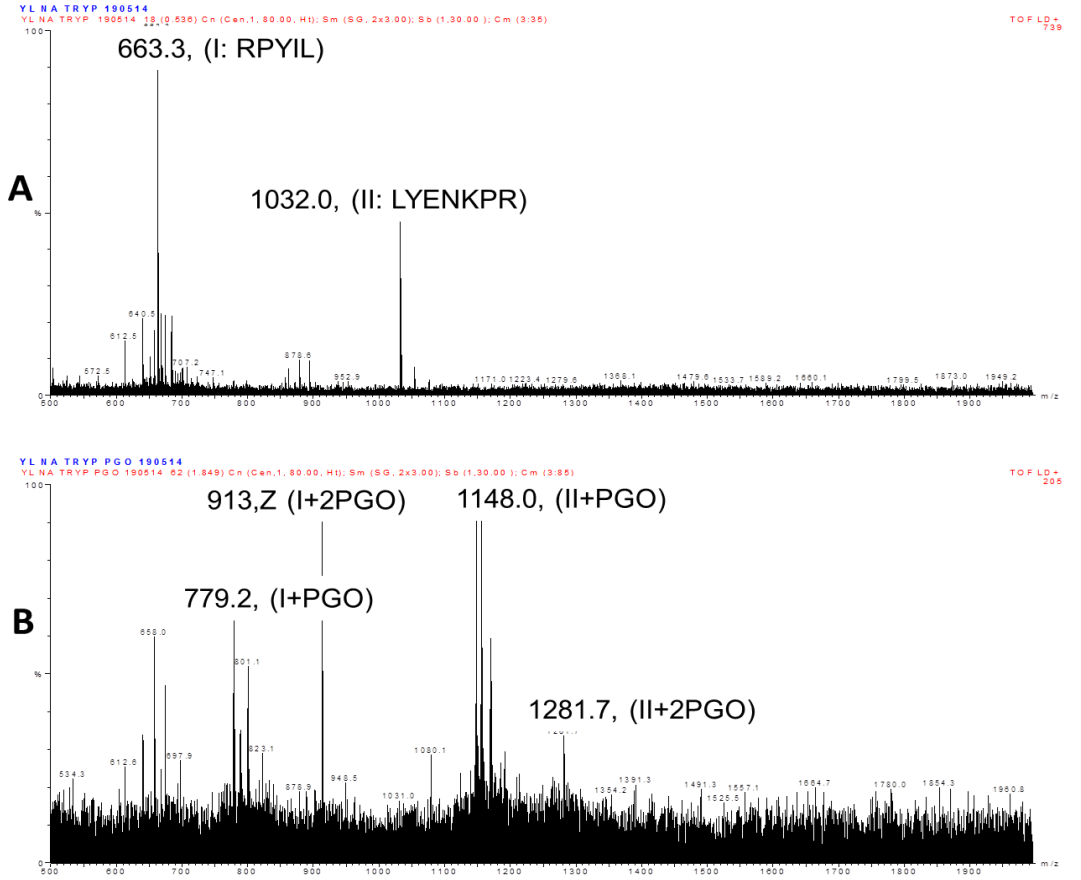


Figure 6.6 Mass spectra of peptide I and II reacted with 200 mM phenylglyoxal for 30 min. A. NA was digested with trypsin to produce two peptides I (RPYIL) and II (LYENKPR). B. I/II+PGO, peptide I/II modified with one PGO; I/II+2xPGO; peptide I/II modified with two PGO.

6.3.2 Method development of arginine labelling by azidophenylglyoxal (APG)

The reaction of azidophenylglyoxal was examined before its use to label arginine residues protected by binding to heparin. The stock of APG (4 M) was adjusted to 20 mM, 50 mM or 100 mM with PBS. Owing to the greater hydrophobicity of APG compared to PG, it was found that 40 % (v/v) DMSO was required to maintain the highest concentration (200 mM) in solution. Reaction of APG at these different concentrations with peptide FA was then performed for 30 min in the dark. Two major peaks were observed upon MALDI-TOF analysis of the products: unmodified peptide FA and peptide FA modified with one APG (Figs 6.11 A, B and C). After calculating the mass increase following modification of peptide FA, it would appear that the guanidyl group of arginine primarily reacts with the dicarbonyl part of phenylglyoxal to form a Schiff's base complex and this is then followed by a ring expansion formed due to the nitrene group reacting with double bonds, which normally happens after the photoactivation of APG with UV light [217] (Fig. 6.10). When peptide FA was reacted with APG for longer times (50 mM for 2 h), there was still some unmodified peptide FA present (Fig. 6.11D).

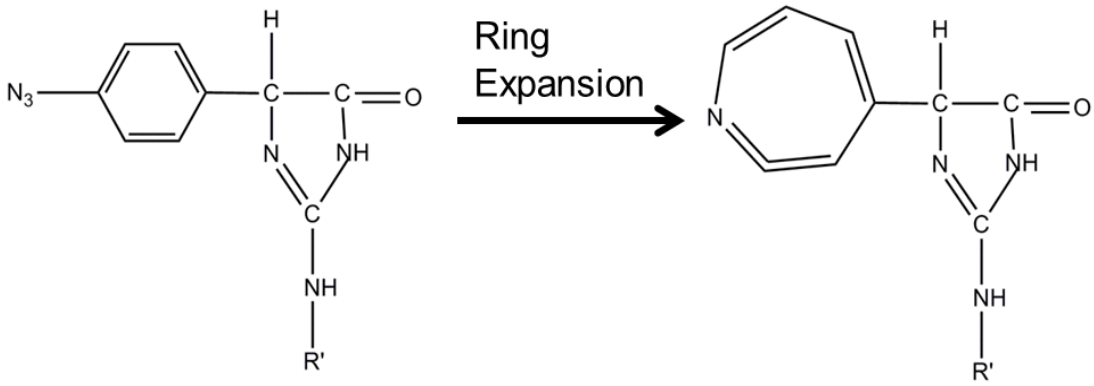


Figure 6.10 Ring expansion caused by the reaction of the nitrene group with double bonds.

When the highest concentration of APG, 200 mM, was used, peptide FA could be modified by one APG or two APG when the reaction was performed for 5 min in the dark, whereas in peptide FA modified with a ratio of 1:1 of APG: arginine, was only found after 30 min reaction (Fig 6.12). However, unmodified peptide FA was identified in both reactions, which means 200 mM APG would not be able to offer full arginine labelling. The arginine labelling of azidophenylglyoxal seems be affected by the reactivity of the azide, in the context of an adjacent aromatic ring.

6.4 Discussion:

To achieve complete labelling, the conditions for modifying arginine with PGO have been optimized. Even though different types of interaction product were identified by mass spectrometry, all arginines were found to be labelled using 200 mM PGO. The lysine targeted labelling technique has simpler and more predictable chemical reactions and is very well established. Thus, it has been used to identify heparin binding sites [170], In contrast, techniques for arginine targeted labelling have not been as well developed and this is the first time it has been explored in the context of identifying heparin binding sites. Although PGO has been used to block arginine for inhibition of enzyme active sites, arginine labelling with PGO has not been developed. This may be due to there being more than one reaction product, which would have been difficult to distinguish in the past. However, the different products are easily identified by mass spectrometry, so this approach can certainly be used in the context of omics experiments.

When APG was used in reactions with arginine, full labelling could not be achieved, even at the highest concentration of APG. One major problem was that the azide adjacent to the aromatic ring caused a ring expansion reaction, which may reduce the solubility of APG and affect its interaction with arginine. Therefore, for further development of this method, APG should be first reacted with DIBO-biotin to form the triazole structure. This will then prevent the azide reacting intramolecularly with this neighbouring aromatic ring. Alternatively, the APG could be replaced by PGO with a different reactive group. For example, a thiol group is able to react with biotin-maleimide. Though maleimide would also interact with cysteine residues, these could be blocked with iodoacetamide before the labelling step. Another

solution would be to use PGO and mass tagged PGO for completing protection and labelling.

In conclusion, the principle of the arginine targeted protect and label technique has been established. Optimization of the arginine labelling steps remains to be achieved. Since all of the measurements conducted here were performed on peptides, the optimised arginine-directed protect and label would need to be tested on proteins with well-defined heparin binding sites. The FGFs used in Chapter 4 and in Xu *et al*. [34], would provide such a test system to validate the method.

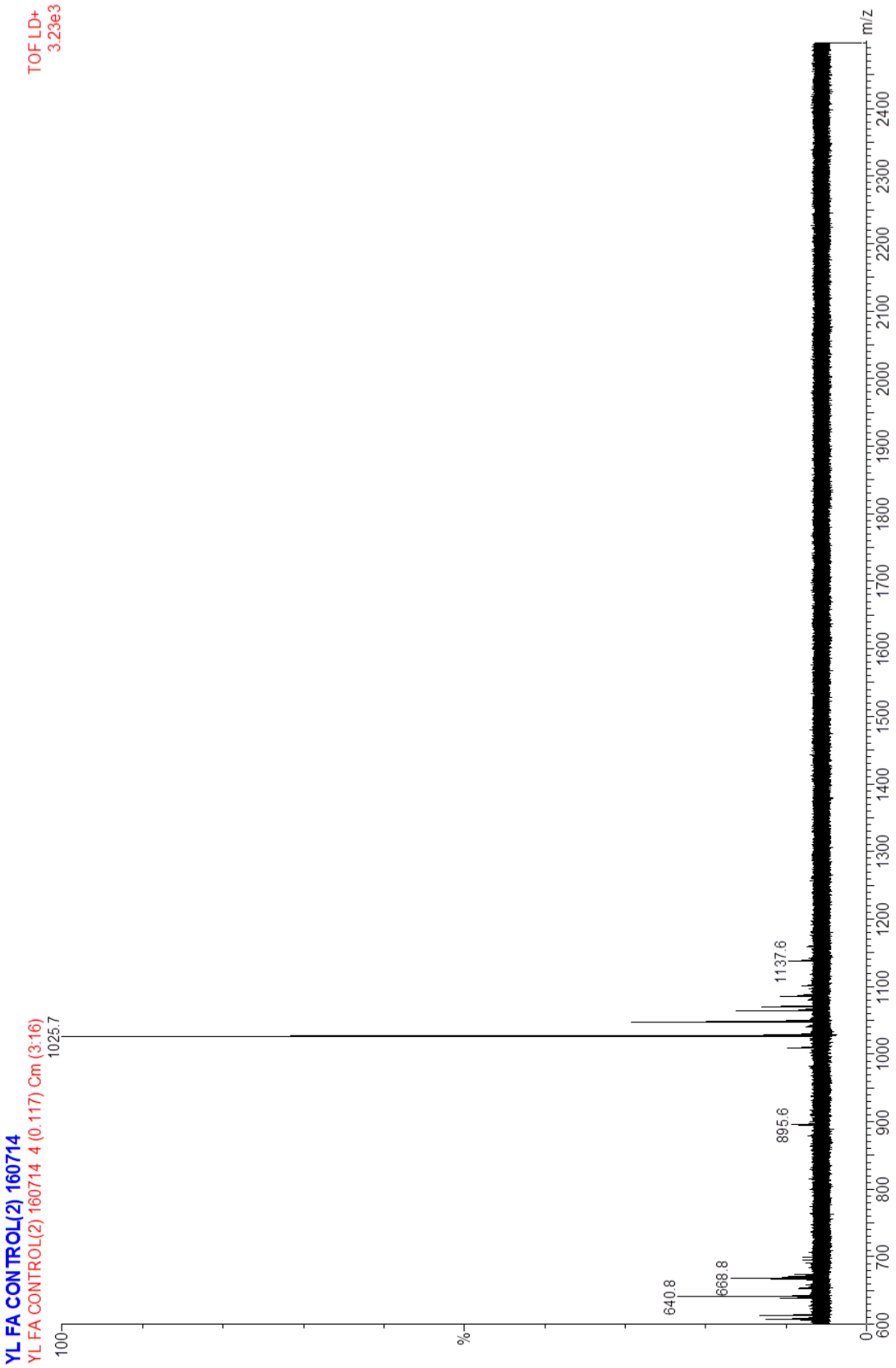


Figure 6.7 Mass spectra of FA. FA (1.5 μg) was also loaded on MALDI-TOF as mass control. The mass of FA is 1025.7.

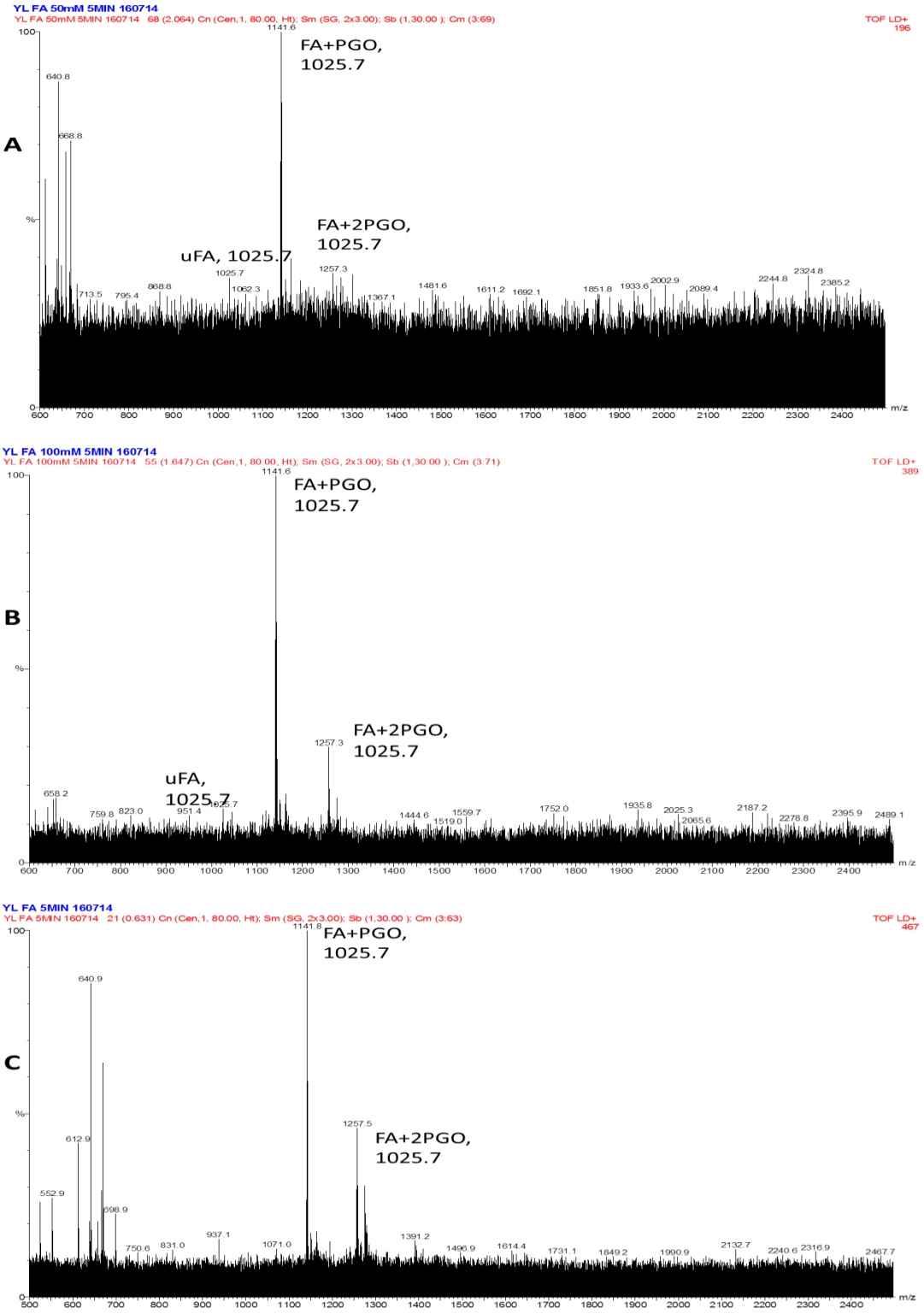


Figure 6.8 Mass spectra of FA reacted with PGO. Peptide FA was reacted with phenylglyoxal (A. 50 mM; B. 100 mM; C. 200 mM) for 5 min. FA, Fibronectin adhesion-promoting peptide; uFA, FA alone; FA+PGO, peptide FA modified with one PGO; FA+2xPGO; peptide FA modified with two PGO.

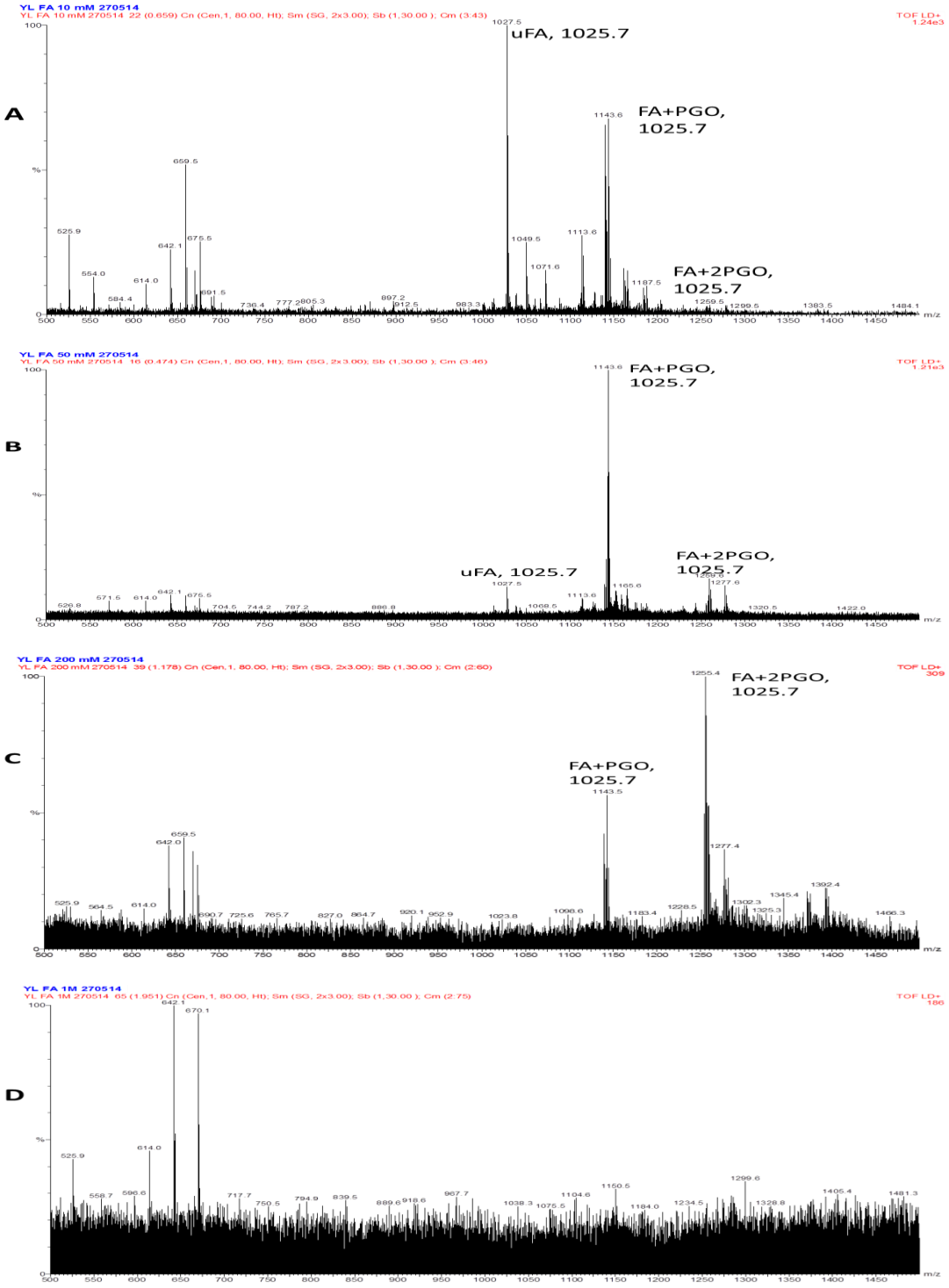


Figure 6.9 Mass spectra of FA reacted with PGO. Peptide FA was reacted with phenylglyoxal (A. 10 mM; B. 50 mM; C. 200 mM; D. 1 M) for 30 min. FA, Fibronectin adhesion-promoting peptide; uFA, FA alone; FA+PGO, peptide FA modified with one PGO; FA+2xPGO; peptide FA modified with two PGO. D. No expected peptide was detected.

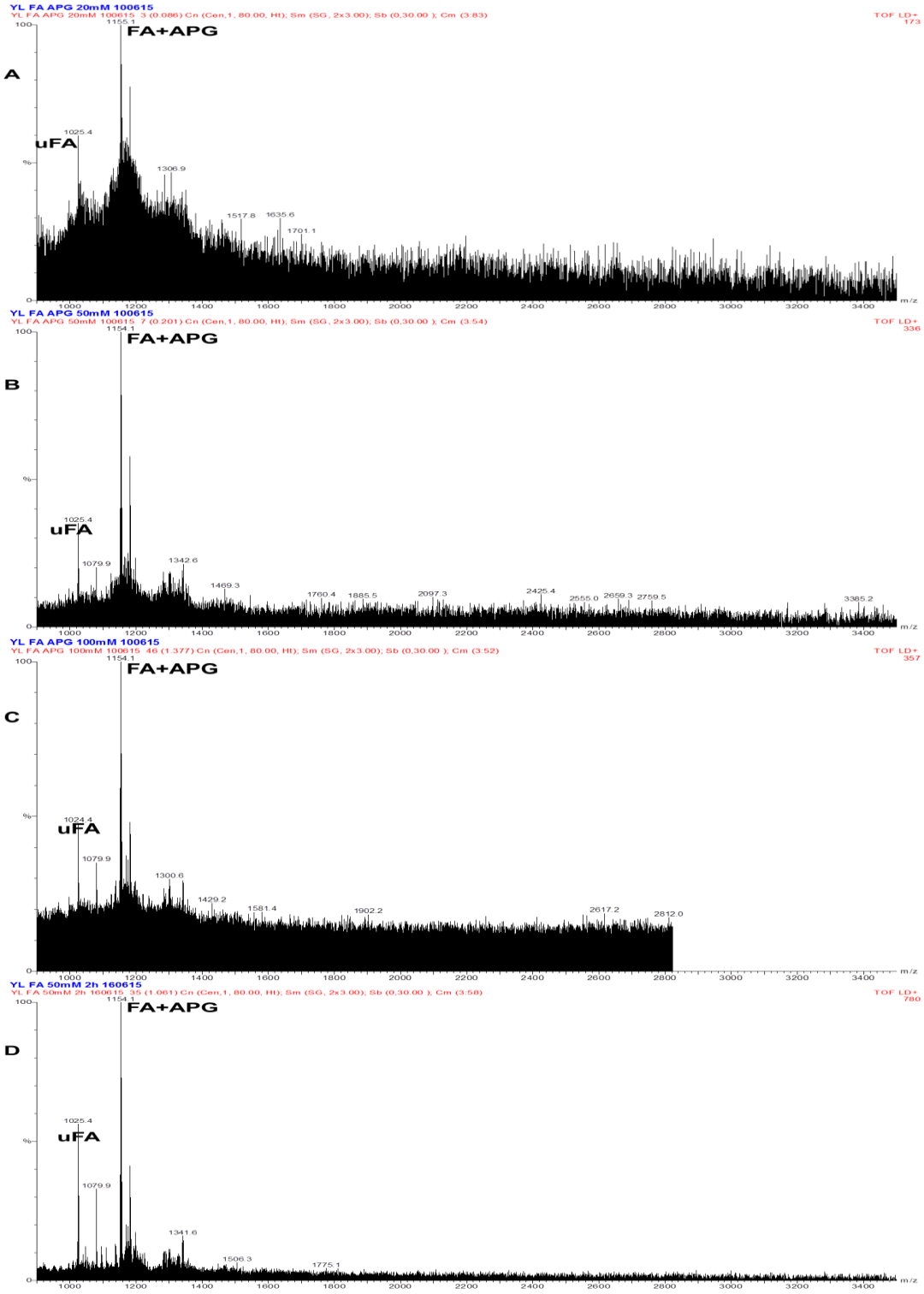


Figure 6.11 Mass spectra of FA reacted with APG. Peptide FA was modified with *p*-azidophenylglyoxal (A. 20 mM; B. 50 mM; C. 100 mM) for 30 min. FA, Fibronectin adhesion-promoting peptide; uFA, FA alone; D. FA+APG, peptide FA modified with one APG.

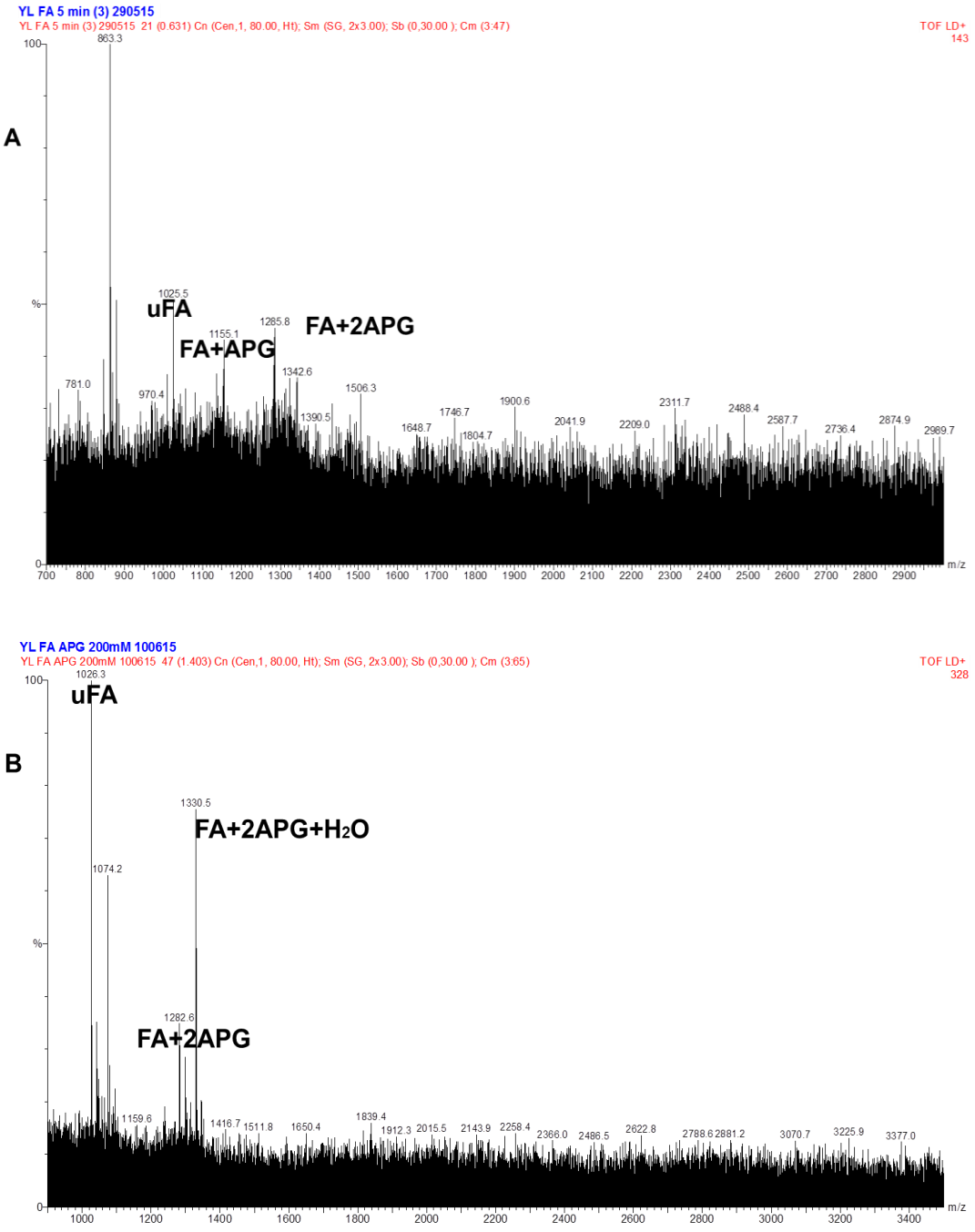


Figure 6.12 Mass spectrum of modified FA by APG. Peptide FA was modified with 200 mM p-azidophenylglyoxal for 5 min (A.) and 30 min (B.) in the dark. FA, Fibronectin adhesion-promoting peptide; uFA, unmodified FA; FA+APG, peptide FA modified with one APG; FA+2APG/2APG+H_2O, peptide FA modified with two APG or two APG plus one H_2O.

Chapter 7 General discussion and perspective

7.1 Discussion:

A large number (>883) of extracellular regulatory proteins have been identified as heparin binding proteins by virtue of their interaction with heparin [113, 114]. For many of these, but not all, their physiological activities will be regulated by HS. FGFs often serve as a model for understanding the consequences of the interaction with HS of other proteins. The expansion of FGF and FGFR from two or three FGF ligands and one *fgfr* alongside HS in the fly *Drosophila* and the worm *C. elegans* to 22 FGFs and 5 *fgfr*s in vertebrates and mammals is reflected in the more complex specification required for body structure in mammals. Thus, the expansion and diversification of the FGF communication system is likely to have been subjected to the same natural selection pressures as those that drove the increasing complexity of animal bodies. The specificity of FGF ligands for FGFR isoforms is reflected in the evolutionary relationship of the FGF ligands at the level of their phylogeny deduced from amino acid sequence. However, the level of specificity of FGFs for binding HS is still debated. One school, drawing on the high specificity of the interaction of antithrombin III with for a pentasaccharide structure in heparin, contends that there are rare and unusual sequences of saccharides in HS, and related heparin, e.g., '$GlcN_{NAc/NS,\ 6S}$-GlcA-$GlcN_{NS,\ 3,\ 6S}$-idoUA2s-$GlcN_{NS,\ 6S}$' [163], responsible for high affinity and high specificity binding of the FGFs. This view is somewhat undermined by the existence of antithrombin III oligosaccharide binding structures that are not identical to the pentasaccharide sequence [186, 197]. Another school contends that largely non-specific ion-exchange interactions underpin FGF binding by the sugar.

However, this view is also undermined, because the concentration of NaCl required for elution of FGFs from a heparin column are higher than that required for elution from a strong cation-exchange column [171]. Most of the work on FGFs is not systematic. That is, only one or two FGFs and/or a limited repertoire of sugar structures have been examined. The one exception is the work of Xu *et al.*, [34], in which the heparin-binding properties of six FGFs from five subfamilies were studied. However, one of these FGFs was FGF21, which is an endocrine member of the family and, thus, does not bind heparin. Moreover, only for the FGF1 subfamily was there more than a single FGF from a subfamily used in this study. Therefore, although the conclusion of this work was that the binding of FGFs to heparin-derived structures was selective and had followed the evolutionary divergence of the FGF family, this was very much a hypothesis that required testing. Therefore, in this thesis, the evolutionary relationship of the FGF family was used as a defined system to determine the specificity of interactions of FGFs with heparin/HS and other model GAGs. Six recombinant FGFs (FGF3, FGF4, FGF6, FGF10, FGF17, and FGF20) from 4 subfamilies were produced (Chapter 3) and their interactions with GAGs probed. To begin to characterize the functions of the secondary heparin binding sites of FGFs, FGF2 with HBS2 or HBS3 mutated to alanine was produced and compared to wild-type FGF2. In addition, an arginine targeted protect and label method was developed, which will enable the identification of HBSs enriched in arginine residues, as these are currently not detectable.

In Chapter 3, in addition to describing the process of the production of FGF proteins, we discovered that N-terminal HaloTag fusions could enhance protein expression and stabilise the FGFs. Thus, comparing production of 11 FGF proteins with an N-terminal HaloTag fusion to those without this, the yield of FGF2, FGF3 and FGF7

was found to have been significantly increased. In addition, HaloTag also was found to markedly enhance the expression as soluble proteins of FGFs that had hitherto been expressed mostly in an insoluble form (FGF3, FGF6, FGF7, FGF8, FGF16, FGF17, FGF20, and FGF22). This was a major step forward, particularly in the cases of FGF6, FGF8, FGF16, and FGF20 that had previously been reported to be expressed mainly in inclusion bodies, and even as truncated proteins [218-221]. Most interestingly, HaloTag was able to prevent the toxicity of FGF7, which usually requires strong repression, e.g., the use of pLysS, that in turn reduces protein yield [222]. Another advantage of HaloTag is that, thanks to its low isoelectric point, anion-exchange chromatography could be used to isolate the HaloTag FGFs, which is orthogonal to the usual approaches for FGF purification. Subsequent cleavage with TEV then allowed the efficient removal of the HaloTag. However, some FGFs appear to aggregate after separation from the HaloTag fusion, indicating that these are not stable. There is clearly not a universal solution to resolve problems of protein expression and for those FGFs that aggregated after cleavage of the Halotag, e.g., FGF6, some engineering of the FGF will likely be necessary. The FGFs that express as insoluble protein all have long N- and/or C-terminal extensions. An NMR analysis of the N-terminus of FGF2 that extends beyond the first beta strand of the beta trefoil has shown that it is very flexible [223]. Therefore, an analysis of the relative disorder of these regions and judicious trimming may yield protein that does not aggregate after cleavage of an expression tag such as HaloTag.

In Chapter 4, the specificity of interaction with the polysaccharide of six FGFs from four subfamilies has been explored at different levels: the heparin structures required for binding by FGFs were measured by differential scanning fluorimetry (DSF) and the heparin binding sites on FGFs were identified by ‘protect and label’. Along with

our existing work on six other FGFs from five FGF subfamilies, there are now at least two members of each paracrine FGF subfamily that have been characterized in this way. This provides a reasonable coverage of the FGF family (Fig. 1.2). When these data are mapped to the FGF phylogenetic tree based on amino acid sequence it becomes clear that FGFs from the same subfamily prefer similar sulfation patterns and lengths of oligosaccharide, and share similar heparin binding sites compared to FGFs from different subfamilies (Fig. 13, Chapter 4). Therefore, it is likely that these HS binding properties of the paracrine FGFs were selected for during the divergence and expansion of the FGF family that was associated with the evolution of more complex animal body plans. It is interesting to note that the specificity of FGFs for FGFR isoforms seems also to be associated with the same evolutionary process (Table 1.2). With respect to HS binding, it seems clear that there is indeed a good degree of specificity and selectivity, but this is not a simple one to one code. The same is true for the interactions of FGFs with FGFR isoforms [35, 51, 169]. It is also intriguing to note that there exist differences between members of a subfamily in terms of their interactions with the polysaccharide, e.g., the likely presence of a secondary HBS3 in FGF20, compared to FGF9 that only has a single, enlarged HBS1.

The discovery that the molecular basis of the interaction of FGFs with GAGs has been subject to the natural selection process that gave rise to an expanded FGF family has important ramifications. One relates to the question of how specific and selective protein-GAG interactions are. There are clearly limitations in describing the molecular basis of these interactions on the basis of the linear structure of the polysaccharide. It is the spatial positions of groups on the sugar that are important for binding and the observation that polysaccharides unrelated to HS can bind and

also participate in the ternary signalling complex with an FGFR [196] provides strong evidence for this. The sugar chain will adopt a variety of conformations in solution and pendant sulfate groups will modify the conformational space that the chain can occupy, which has been demonstrated by NMR and CD studies [224]. In addition, the coordination of cations modifies the conformation of the polysaccharide chain [225]. Finally, while the binding to polysaccharide clearly changes the conformation of the protein, e.g., thermal stabilisation observed by DSF, the reverse is also true: binding to protein alters the conformation of the polysaccharide. The latter point is elegantly made by the co-crystal structure of FGF2 and a heparin dp 6, in which the latter has iduronate residues in both the $^{1}C_{4}$ and the $^{2}S_{0}$ configurations [40]. Thus, the selectivity and specificity identified here is somewhat artificial, since the conformation of HS *in vivo* in extracellular and pericellular matrix will depend on the sequence of saccharides, the coordinated cations and the pre-existing interactions of the HS chain with endogenous proteins. At some level, cells can sense what functional structures they produce and modify these. This is shown by the HS 2-O sulfotransferase knockout mouse, which dies at birth due to kidney agenesis [226]. HS or heparin lacking 2-O-sulfate cannot bind FGF2 or form a productive ternary complex with the FGFR [125]. Yet, the knockout mice have no FGF2 phenotype. Moreover, when embryonic fibroblasts were derived from these mice, their HS did not possess any 2-O-sulfated HS, but the HS was capable of interacting with FGF2 and enabling it to bind and activate FGFR on cells [227]. Thus, there are homeostatic mechanisms whereby cells can modify the chains they produce and perhaps cations coordinated to HS and/or the endogenous proteins bound to this HS, to ensure that as many as possible of the appropriate functions are maintained after perturbation. Such homeostatic plasticity can be considered to be advantageous, since it provides for a

robust, rather than a brittle communication system. However, this clearly limits the degree to which one can apply simple interpretations to the molecular basis of specificity of FGF-GAG interactions. These considerations are likely to apply to many, if not all such interactions, since even the pentasaccharide structure that binds antithrombin III which is used as an example of a highly specific interaction has alternatives [163].

Another ramification of the interaction of FGFs with GAGs being subject to natural selection relates to the phylogenetic relationship of the FGFs themselves. In Chapter 1 (Sections 1.2 and 1.3) two different phylogenetic trees were presented. One is based on the idea that genes that are close to each other are likely to reflect a conservation of an ancestral genome organization; that is, this is less likely to have arisen by chance. The other is based on amino acid sequence homology and again this is deemed to reflect a common ancestor, where amino acids important for structure and function are considered less likely to change over time and their conservation is thus not due to chance. The differences in the two trees relate to the positions of FGF3 and FGF5; both trees cannot be correct and one is thus more likely to reflect a chance event than the other. The results presented here indicate that the phylogeny based on amino acid sequence also reflects conservation of secondary GAG-binding sites in the FGFs and preferences for particular binding structures in GAGs. Although there are no data on FGF5, the data on FGF3 suggest that it is more closely related to the FGF7 subfamily than the FGF4 subfamily, since it has a HBS4 (only present in FGF7 subfamily) and its preferred binding structures in the model GAGs are most similar to FGF7 and FGF10. In addition, work on the preferences of FGFs for FGFR isoforms also show a clear relationship between the amino acid based phylogeny and FGFR binding. FGF3 and FGF5 are clearly more closely

related to the FGF7 and FGF4 subfamilies, respectively, in terms of FGFR isoform preference; for example, FGF5, like FGF4 and FGF6, does not bind FGFR1b, whereas FGF1 and FGF2 do [35, 51]. The function of FGFs is determined by binding to GAGs and FGFR. Natural selection operates on phenotype; it is a particular phenotype that has some advantage over others. The data on the HS binding preferences of the FGFs, coupled to their FGFR isoform preferences are consistent with the phylogeny based on amino acid sequence homology. This suggests that this phylogeny may have been less likely to have arisen by chance than that based on genome organization.

The preference of FGFs from different sub-families for distinct sugar structures will be an important means for cells to regulate the activities of multiple FGFs, independently of each other, in the same tissue. One level at which such regulation can occur is at that of the formation of the ternary FGF:HS:FGFR signaling complex (Section 1.7). Clearly, if a cell produces HS structures that do not bind members of a particular FGF subfamily or an individual FGF, then even if the appropriate (Table 1.2) FGFR isoform is expressed on that cell, a growth response cannot occur, because a ternary complex cannot be formed. Another level of regulation relates to the diffusion of FGFR ligands from a source to a target cell. The binding sites in the pericellular matrix of Rama 27 fibroblasts were shown to control the diffusion of FGF2 [178]. An important question raised by this work was whether binding to HS might control the diffusion of HS-binding ligands more generally. This question has recently been addressed by measurements on the binding and diffusion of five Halo-FGFs (Halo-FGF1, Halo-FGF2, Halo-FGF6, Halo-FGF10, Halo-FGF20) in the pericellular matrix of fixed Rama 27 fibroblasts [228]. Fixed cells were used to ensure there were no confounding effects from the membrane and cell movement. The data

clearly demonstrate that the binding preferences of FGFs for distinct sugar structures determine both their interactions with glycosaminoglycans in the pericellular matrix and their diffusion [228]. For example, more Halo-FGF1 bound to the pericellular matrix than Halo-FGF2 or Halo-FGF6, which is consistent with FGF1 binding a broader range of structures in HS (Figure 13, Chapter 4). Moreover, Halo-FGF20, which has very unusual (for an FGF) binding preferences (Figure 13, Chapter 4) failed to bind detectably. FGF10 was found to be bound to both HS and CS/DS. Interestingly, the expression of binding structures for the different Halo-FGFs showed spatial heterogeneity within the pericellular matrix of individual cells, which was most pronounced for Halo-FGF10, which bound in patches, separated by areas where Halo-FGF10 did not bind. The diffusion of the FGFs that bound to the pericellular matrix was also found to be different, with Halo-FGF2 and Halo-FGF6 diffusing faster than Halo-FGF1. This is likely to be due to Halo-FGF1 possessing more binding sites in the pericellular matrix of these cells [229]. Thus, cells express HS chains with a range of structures that can selectively bind different FGFs, and the HS chains with particular FGF binding properties are segregated, that is they can be localized differently in the pericellular matrix. The regulation of diffusion of FGFs by binding to different structures in HS that are distributed heterogeneously in the pericellular matrix allow cells to control independently the localization and the bioavailability of the multipole FGFs simultaneously.

In Chapter 5, the mutants of FGF2 (HBS2 and HBS3), in which either HBS2 or HBS3 of FGF2 were mutated by replacing the lysines and arginines with alanines, have been produced to explore the secondary HBSs and its function. Both mutants have been compared to wild-type FGF2 in terms of the heparin structures required

for optimum binding and their biological activities on Rama 27 fibroblasts. The mutants were found to have a lower stability than wild-type FGF2, which may be caused by the hydrophobic pockets formed by the alanines. The mutants were found to have similar preference for sulfation patterns. However, differences between wild-type FGF2 and the mutants were found in their interactions to polysaccharides with a low level of sulfation and of restricted length. Compared to wild-type FGF2, FGF2 (HBS2) binds N-sulfated heparin strongly and both mutants required larger structures for minimal binding, which could be due to a reduction of their rebinding capacity to the polysaccharide after (partial) dissociation. FGF2 and its mutants have been shown to stimulate similarly the phosphorylation of $p42/44^{MAPK}$ and cell proliferation, which indicates that the HBS2 and HBS3 of FGF2 may not be directly involved in these aspects of FGF signalling. Since the HBSs have been suggested to contribute to the cross-linking of HS chains observed *in vitro* [198], their role may instead be to control the diffusion of FGF2 in extracellular matrix, and be responsible for at least some of the confined motion observed previously [178].

In Chapter 6, a new arginine targeted 'protect and label' technique was developed to allow a complete identification of the residues in HBSs involved in ionic bonding. In addition, there are some HBSs that only contain arginine and these would then be revealed. A first challenge was to distinguish the different reaction products of arginine and PGO. This difficulty is likely to underlie the lack of work exploiting this chemistry for the identification of arginine residues in binding sites of proteins and for labelling proteins. The present work demonstrates that the different reaction products are predictable in terms of the position and neighbours of the arginine residues and resolvable by mass spectrometry. The result is that the different reaction

products should, when the method is applied to proteins, provide additional confirmatory information relating to the arginines that are identified.

However, 4-azidophenylglyoxal (APG), which was used in the labelling step, did not react stoichiometrically with arginine residues. This seems to be due to the interaction of arginine and APG and the unforeseen side reaction of ring expansion reaction due to the azide group adjacent to the aromatic ring. Therefore, the labelling step needs to be reconsidered to avoid this side reaction.

7.2 Further work:

The work done for this thesis highlights a number of areas for future work. These relate to the molecular basis of the interaction of FGFs with HS, the functions of the secondary binding sites and the development of the 'protect and label' approach for identifying heparin binding sites in proteins.

Several aspects of the interaction of the six FGFs used in the present work with HS need to be determined. Firstly the binding parameters (K_D, and if possible the kinetic parameters ka and kd). This was attempted, but these proteins exhibited strong non-specific binding to surfaces, which precluded using surface-based optical biosensors. Solution based methods such as microscale thermophoresis [34, 230] may enable such measurements, although preliminary experiments were not conclusive. Alternatives might be to use more robust surfaces for the optical techniques, such as the 2-D semi-crystalline arrays of streptavidin assembled on a biotinylated support grafted onto gold surfaces [198]. Circular dichroism would provide insight into the proportions of secondary structure present in these FGFs in solution and the degree

to which this changes upon binding to heparin. In previous work [34] it was shown that the FGFs investigated contained less secondary structure than expected from their crystal structure, but that this increased when they bound heparin. The DSF data support this, and acquiring direct evidence would lend weight to the idea that polysaccharide-induced conformational change in the FGF is important for FGFR binding [119], which may, therefore, be worth revisiting. Clearly, the development of an arginine-targeted protect and label method is important and applicable beyond the realm of FGFs. The first step would be to optimize the arginine labelling with APG. One way to avoid the ring expansion problem would be to first react the APG with DIBO-biotin in a so-called "click" reaction. This would remove the azide, and should prevent ring expansion. Alternatives would be to use a PGO with a different reactive group, such as a thiol group, which is able to react with biotin-maleimide or to simply use a mass tagged PGO.

In the longer term, there are two areas that thesis suggests would be useful to investigate. These relate to the secondary HBS of FGFs, and to the identification of heparin binding proteins in the extracellular matrix and understanding how these interactions may change either due to physiological challenge or to disease.

For the characterization of the secondary binding sites of FGFs, the HBS mutations need to be redesigned by replacing the lysines and arginines of HBS with amino acids identified by alignment with endocrine FGFs, which should avoid the low stability of the FGF2 mutants produced in this work. Based on work with FGF2 and FGF9, it has been proposed that the secondary HBSs of FGFs are able to cross-link HS [198]; this idea could be examined in depth using a combination of FGFs with different secondary HBSs and mutants of these. This could then be taken a step

further, by measuring the movement of these FGFs (wild-type and mutants) in cell pericellular matrix, e.g., Duchesene et al., 2012, to determine whether the ability to cross link HS chains is associated with confined motion of FGF.

With a 'protect and label' procedure targeting lysines and arginines it will be possible to identify the major ionic contacts in the HBS of proteins. However, for proteins with multiple HBSs, it does not identify, which of these is engaged *in vivo* and whether this might change in the course of a physiological challenge or in disease. A proteomics pipeline has been developed to identify heparin-binding proteins from tissues [114, 156]. However the weakness of this approach is that it cannot distinguish a heparin binding protein from a protein bound to a heparin binding protein. In addition, it provides no information on which HBSs are engaged with the polysaccharide. This pipeline uses heparin affinity chromatography to purify heparin binding proteins followed by their identification using mass spectrometry. It should be possible to integrate the 'protect and label' strategy into this the pipeline. The protection step would be performed on purified extracellular matrix and following dissociation of protein-polysaccharide complexes with high NaCl, exposed lysines or arginines could then be labelled. While the conditions used to dissociate the protein-polysaccharide complexes are likely to require substantial optimization, integration of the methods would have several benefits. Firstly, the HBSs actually engaged with the polysaccharide would be identified; comparison with the proteins isolated by heparin affinity chromatography would then identify proteins bound to the heparin-binding proteins. Secondly, a comparison of normal versus disease, e.g., pancreas and acute pancreatitis [114], would yield information not

just on changes of levels of proteins, but also changes in their association with GAGs, which would provide a much richer insight into the disease process.

Supplemental data

Papers and manuscripts

Contributions to work

Paper 1

D. J. Nieves, **Y. Li**, D. G. Fernig, R. Lévy, 'Photothermal raster image correlation spectroscopy of gold nanoparticles in solution and on live cells', Open Science, 2 (2015).

Produced recombinant FGF-2 protein and contributed to edit paper.

Paper 2

E. Migliorini, D. Thakar, J. Kuhnle, R. Sadir, D. P. Dyer, **Y. Li**, C. Sun, B. F. Volkman, T. M. Handel, L. Coche-Guerente, D. G. Fernig, H. Lortat-Jacob, and R. P. Richter, 'Cytokines and Growth Factors Cross-Link Heparan Sulfate', Open Biol, 5 (2015).

Produced recombinant FGF-2 protein and FGF-9 protein and contributed to edit paper.

Paper 3

C. Sun, **Y. Li**, S. E. Taylor, X. Mao, M. C. Wilkinson, and D. G. Fernig, 'Halotag Is an Effective Expression and Solubilisation Fusion Partner for a Range of Fibroblast Growth Factors', PeerJ, 3 (2015).

Contributed to the cloning of cDNA of Histag-FGF3, Histag-FGF7, Halotag/Histag-FGF16 and Halotag/Histag-17. Expression and Production of FGFs. Assisted with biological activities assay. Data analysis. **Co-wrote** the paper.

Paper 4

Q. M. Nunes, **Y. Li**, C. Sun, T. K. Kinnunen, D. G. Fernig, 'Fibroblast growth factors as tissue repair and regeneration therapeutics', PeerJ, accepted (2015).

Contributed to the translation of Chinese papers. **Co-wrote** the paper.

Paper 5

Y. Li, C. Sun, E. A. Yates, M. C. Wilkinson, and D. G. Fernig, 'Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification, Open Biol, Submited (2015).

Produced recombinant FGF proteins, performed the experiments and drafted the entire manuscript. Other authors supplied sugars and FGF proteins.

References:

1. Gospodarowicz, D., *Localisation of a fibroblast growth factor and its effect alone and with hydrocortisone on 3T3 cell growth.* Nature, 1974. **249**(453): p. 123-7.
2. Gospodarowicz, D., *Purification of a fibroblast growth factor from bovine pituitary.* J Biol Chem, 1975. **250**(7): p. 2515-20.
3. Gospodarowicz, D., H. Bialecki, and G. Greenburg, *Purification of the fibroblast growth factor activity from bovine brain.* J Biol Chem, 1978. **253**(10): p. 3736-43.
4. Maciag, T., et al., *An endothelial cell growth factor from bovine hypothalamus: identification and partial characterization.* Proc Natl Acad Sci U S A, 1979. **76**(11): p. 5674-8.
5. Dickson, C., et al., *Expression, processing, and properties of int-2.* Ann N Y Acad Sci, 1991. **638**: p. 18-26.
6. Yoshida, T., et al., *Characterization of the hst-1 gene and its product.* Ann N Y Acad Sci, 1991. **638**: p. 27-37.
7. Goldfarb, M., et al., *Expression and possible functions of the FGF-5 gene.* Ann N Y Acad Sci, 1991. **638**: p. 38-52.
8. Coulier, F., et al., *The FGF6 gene within the FGF multigene family.* Ann N Y Acad Sci, 1991. **638**: p. 53-61.
9. Aaronson, S.A., et al., *Keratinocyte growth factor. A fibroblast growth factor family member with unusual target cell specificity.* Ann N Y Acad Sci, 1991. **638**: p. 62-77.
10. Tanaka, A., et al., *Cloning and characterization of an androgen-induced growth factor essential for the androgen-dependent growth of mouse mammary carcinoma cells.* Proc Natl Acad Sci U S A, 1992. **89**(19): p. 8928-32.
11. Miyamoto, M., et al., *Molecular cloning of a novel cytokine cDNA encoding the ninth member of the fibroblast growth factor family, which has a unique secretion property.* Mol Cell Biol, 1993. **13**(7): p. 4251-9.
12. Yamasaki, M., et al., *Structure and expression of the rat mRNA encoding a novel member of the fibroblast growth factor family.* J Biol Chem, 1996. **271**(27): p. 15918-21.
13. Smallwood, P.M., et al., *Fibroblast growth factor (FGF) homologous factors: new members of the FGF family implicated in nervous system development.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9850-7.
14. Miyake, A., et al., *Structure and expression of a novel member, FGF-16, on the fibroblast growth factor family.* Biochem Biophys Res Commun, 1998. **243**(1): p. 148-52.
15. Hoshikawa, M., et al., *Structure and expression of a novel fibroblast growth factor, FGF-17, preferentially expressed in the embryonic brain.* Biochem Biophys Res Commun, 1998. **244**(1): p. 187-91.
16. Ohbayashi, N., et al., *Structure and expression of the mRNA encoding a novel fibroblast growth factor, FGF-18.* J Biol Chem, 1998. **273**(29): p. 18161-4.
17. Ohmachi, S., et al., *FGF-20, a novel neurotrophic factor, preferentially expressed in the substantia nigra pars compacta of rat brain.* Biochem Biophys Res Commun, 2000. **277**(2): p. 355-60.
18. Nishimura, T., et al., *Identification of a novel FGF, FGF-21, preferentially expressed in the liver.* Biochim Biophys Acta, 2000. **1492**(1): p. 203-6.
19. Nishimura, T., et al., *Structure and expression of a novel human FGF, FGF-19, expressed in the fetal brain.* Biochim Biophys Acta, 1999. **1444**(1): p. 148-51.

20. Nakatake, Y., et al., *Identification of a novel fibroblast growth factor, FGF-22, preferentially expressed in the inner root sheath of the hair follicle.* Biochim Biophys Acta, 2001. **1517**(3): p. 460-3.
21. Yamashita, T., M. Yoshioka, and N. Itoh, *Identification of a novel fibroblast growth factor, FGF-23, preferentially expressed in the ventrolateral thalamic nucleus of the brain.* Biochem Biophys Res Commun, 2000. **277**(2): p. 494-8.
22. Ornitz, D.M. and N. Itoh, *Fibroblast growth factors.* Genome Biol, 2001. **2**(3): p. REVIEWS3005.
23. Sutherland, D., C. Samakovlis, and M.A. Krasnow, *branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching.* Cell, 1996. **87**(6): p. 1091-101.
24. Burdine, R.D., et al., *egl-17 encodes an invertebrate fibroblast growth factor family member required specifically for sex myoblast migration in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2433-7.
25. Roubin, R., et al., *let-756, a C. elegans fgf essential for worm development.* Oncogene, 1999. **18**(48): p. 6741-7.
26. Stathopoulos, A., et al., *pyramus and thisbe: FGF genes that pattern the mesoderm of Drosophila embryos.* Genes Dev, 2004. **18**(6): p. 687-99.
27. Gryzik, T. and H.A. Muller, *FGF8-like1 and FGF8-like2 encode putative ligands of the FGF receptor Htl and are required for mesoderm migration in the Drosophila gastrula.* Curr Biol, 2004. **14**(8): p. 659-67.
28. Satou, Y., K.S. Imai, and N. Satoh, *Fgf genes in the basal chordate Ciona intestinalis.* Dev Genes Evol, 2002. **212**(9): p. 432-8.
29. Itoh, N., *The Fgf families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease.* Biol Pharm Bull, 2007. **30**(10): p. 1819-25.
30. Itoh, N. and D.M. Ornitz, *Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease.* J Biochem, 2011. **149**(2): p. 121-30.
31. Itoh, N. and D.M. Ornitz, *Evolution of the Fgf and Fgfr gene families.* Trends Genet, 2004. **20**(11): p. 563-9.
32. Horton, A.C., et al., *Phylogenetic analyses alone are insufficient to determine whether genome duplication(s) occurred during early vertebrate evolution.* J Exp Zool B Mol Dev Evol, 2003. **299**(1): p. 41-53.
33. Itoh, N. and D.M. Ornitz, *Functional evolutionary history of the mouse Fgf gene family.* Dev Dyn, 2008. **237**(1): p. 18-27.
34. Xu, R., et al., *Diversification of the structural determinants of fibroblast growth factor-heparin interactions: implications for binding specificity.* J Biol Chem, 2012. **287**(47): p. 40061-73.
35. Ornitz, D.M., et al., *Receptor specificity of the fibroblast growth factor family.* J Biol Chem, 1996. **271**(25): p. 15292-7.
36. Ornitz, D.M., *FGFs, heparan sulfate and FGFRs: complex interactions essential for development.* Bioessays, 2000. **22**(2): p. 108-12.
37. Zhu, X., et al., *Three-dimensional structures of acidic and basic fibroblast growth factors.* Science, 1991. **251**(4989): p. 90-3.
38. Zhang, J.D., et al., *Three-dimensional structure of human basic fibroblast growth factor, a structural homolog of interleukin 1 beta.* Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3446-50.
39. Li, L.Y., et al., *Diminished heparin binding of a basic fibroblast growth factor mutant is associated with reduced receptor binding, mitogenesis, plasminogen activator induction, and in vitro angiogenesis.* Biochemistry, 1994. **33**(36): p. 10999-1007.

40. Faham, S., et al., *Heparin structure and interactions with basic fibroblast growth factor.* Science, 1996. **271**(5252): p. 1116-20.
41. Thompson, L.D., M.W. Pantoliano, and B.A. Springer, *Energetic characterization of the basic fibroblast growth factor-heparin interaction: identification of the heparin binding domain.* Biochemistry, 1994. **33**(13): p. 3831-40.
42. Baird, A., et al., *Receptor- and heparin-binding domains of basic fibroblast growth factor.* Proc Natl Acad Sci U S A, 1988. **85**(7): p. 2324-8.
43. Mohan, S.K., S.G. Rani, and C. Yu, *The heterohexameric complex structure, a component in the non-classical pathway for fibroblast growth factor 1 (FGF1) secretion.* J Biol Chem, 2010. **285**(20): p. 15464-75.
44. Nickel, W., *Pathways of unconventional protein secretion.* Curr Opin Biotechnol, 2010. **21**(5): p. 621-6.
45. Trudel, C., et al., *Translocation of FGF2 to the cell surface without release into conditioned media.* J Cell Physiol, 2000. **185**(2): p. 260-8.
46. Nickel, W., *The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes.* Eur J Biochem, 2003. **270**(10): p. 2109-19.
47. Revest, J.M., L. DeMoerlooze, and C. Dickson, *Fibroblast growth factor 9 secretion is mediated by a non-cleaved amino-terminal signal sequence.* J Biol Chem, 2000. **275**(11): p. 8083-90.
48. Miyakawa, K. and T. Imamura, *Secretion of FGF-16 requires an uncleaved bipartite signal sequence.* J Biol Chem, 2003. **278**(37): p. 35718-24.
49. Goldfarb, M., et al., *Fibroblast growth factor homologous factors control neuronal excitability through modulation of voltage-gated sodium channels.* Neuron, 2007. **55**(3): p. 449-63.
50. Shakkottai, V.G., et al., *FGF14 regulates the intrinsic excitability of cerebellar Purkinje neurons.* Neurobiol Dis, 2009. **33**(1): p. 81-8.
51. Zhang, X., et al., *Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family.* J Biol Chem, 2006. **281**(23): p. 15694-700.
52. Goetz, R., et al., *Molecular insights into the klotho-dependent, endocrine mode of action of fibroblast growth factor 19 subfamily members.* Mol Cell Biol, 2007. **27**(9): p. 3417-28.
53. Kurosu, H., et al., *Regulation of fibroblast growth factor-23 signaling by klotho.* J Biol Chem, 2006. **281**(10): p. 6120-3.
54. Urakawa, I., et al., *Klotho converts canonical FGF receptor into a specific receptor for FGF23.* Nature, 2006. **444**(7120): p. 770-4.
55. Beenken, A. and M. Mohammadi, *The FGF family: biology, pathophysiology and therapy.* Nat Rev Drug Discov, 2009. **8**(3): p. 235-53.
56. Suh, J.M., et al., *Endocrinization of FGF1 produces a neomorphic and potent insulin sensitizer.* Nature, 2014. **513**(7518): p. 436-9.
57. Zhou, M., et al., *Fibroblast growth factor 2 control of vascular tone.* Nat Med, 1998. **4**(2): p. 201-7.
58. Dono, R., et al., *Impaired cerebral cortex development and blood pressure regulation in FGF-2-deficient mice.* EMBO J, 1998. **17**(15): p. 4213-25.
59. Miller, D.L., et al., *Compensation by fibroblast growth factor 1 (FGF1) does not account for the mild phenotypic defects observed in FGF2 null mice.* Mol Cell Biol, 2000. **20**(6): p. 2260-8.
60. Tobe, T., et al., *Targeted disruption of the FGF2 gene does not prevent choroidal neovascularization in a murine model.* Am J Pathol, 1998. **153**(5): p. 1641-6.
61. Ortega, S., et al., *Neuronal defects and delayed wound healing in mice lacking fibroblast growth factor 2.* Proc Natl Acad Sci U S A, 1998. **95**(10): p. 5672-7.

62. Schultz, J.E., et al., *Fibroblast growth factor-2 mediates pressure-induced hypertrophic response.* J Clin Invest, 1999. **104**(6): p. 709-19.
63. Tekin, M., et al., *Homozygous mutations in fibroblast growth factor 3 are associated with a new form of syndromic deafness characterized by inner ear agenesis, microtia, and microdontia.* Am J Hum Genet, 2007. **80**(2): p. 338-44.
64. Sugi, Y., et al., *Fibroblast growth factor (FGF)-4 can induce proliferation of cardiac cushion mesenchymal cells during early valve leaflet formation.* Dev Biol, 2003. **258**(2): p. 252-63.
65. Sun, X., F.V. Mariani, and G.R. Martin, *Functions of FGF signalling from the apical ectodermal ridge in limb development.* Nature, 2002. **418**(6897): p. 501-8.
66. Feldman, B., et al., *Requirement of FGF-4 for postimplantation mouse development.* Science, 1995. **267**(5195): p. 246-9.
67. Hebert, J.M., et al., *FGF5 as a regulator of the hair growth cycle: evidence from targeted and spontaneous mutations.* Cell, 1994. **78**(6): p. 1017-25.
68. Drogemuller, C., et al., *Mutations within the FGF5 gene are associated with hair length in cats.* Anim Genet, 2007. **38**(3): p. 218-21.
69. Housley, D.J. and P.J. Venta, *The long and the short of it: evidence that FGF5 is a major determinant of canine 'hair'-itability.* Anim Genet, 2006. **37**(4): p. 309-15.
70. Armand, A.S., I. Laziz, and C. Chanoine, *FGF6 in myogenesis.* Biochim Biophys Acta, 2006. **1763**(8): p. 773-8.
71. Floss, T., H.H. Arnold, and T. Braun, *A role for FGF-6 in skeletal muscle regeneration.* Genes Dev, 1997. **11**(16): p. 2040-51.
72. Guo, L., L. Degenstein, and E. Fuchs, *Keratinocyte growth factor is required for hair development but not for wound healing.* Genes Dev, 1996. **10**(2): p. 165-75.
73. Qiao, J., et al., *FGF-7 modulates ureteric bud growth and nephron number in the developing kidney.* Development, 1999. **126**(3): p. 547-54.
74. Liu, A. and A.L. Joyner, *Early anterior/posterior patterning of the midbrain and cerebellum.* Annu Rev Neurosci, 2001. **24**: p. 869-96.
75. O'Leary, D.D., S.J. Chou, and S. Sahara, *Area patterning of the mammalian cortex.* Neuron, 2007. **56**(2): p. 252-69.
76. Meyers, E.N., M. Lewandoski, and G.R. Martin, *An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination.* Nat Genet, 1998. **18**(2): p. 136-41.
77. !!! INVALID CITATION !!!
78. Colvin, J.S., et al., *Lung hypoplasia and neonatal death in Fgf9-null mice identify this gene as an essential regulator of lung mesenchyme.* Development, 2001. **128**(11): p. 2095-106.
79. Kato, S. and K. Sekine, *FGF-FGFR signaling in vertebrate organogenesis.* Cell Mol Biol (Noisy-le-grand), 1999. **45**(5): p. 631-8.
80. Lu, S.Y., et al., *FGF-16 is released from neonatal cardiac myocytes and alters growth-related signaling: a possible role in postnatal development.* Am J Physiol Cell Physiol, 2008. **294**(5): p. C1242-9.
81. Xu, J., Z. Liu, and D.M. Ornitz, *Temporal and spatial gradients of Fgf8 and Fgf17 regulate proliferation and differentiation of midline cerebellar structures.* Development, 2000. **127**(9): p. 1833-43.
82. Liu, Z., et al., *Coordination of chondrogenesis and osteogenesis by fibroblast growth factor 18.* Genes Dev, 2002. **16**(7): p. 859-69.
83. Ohbayashi, N., et al., *FGF18 is required for normal cell proliferation and differentiation during osteogenesis and chondrogenesis.* Genes Dev, 2002. **16**(7): p. 870-9.

84. Inagaki, T., et al., *Fibroblast growth factor 15 functions as an enterohepatic signal to regulate bile acid homeostasis.* Cell Metab, 2005. **2**(4): p. 217-25.
85. Fu, L., et al., *Fibroblast growth factor 19 increases metabolic rate and reverses dietary and leptin-deficient diabetes.* Endocrinology, 2004. **145**(6): p. 2594-603.
86. Tomlinson, E., et al., *Transgenic mice expressing human fibroblast growth factor-19 display increased metabolic rate and decreased adiposity.* Endocrinology, 2002. **143**(5): p. 1741-7.
87. Holt, J.A., et al., *Definition of a novel growth factor-dependent signal cascade for the suppression of bile acid biosynthesis.* Genes Dev, 2003. **17**(13): p. 1581-91.
88. Lundasen, T., et al., *Circulating intestinal fibroblast growth factor 19 has a pronounced diurnal variation and modulates hepatic bile acid synthesis in man.* J Intern Med, 2006. **260**(6): p. 530-6.
89. Ohmachi, S., et al., *Preferential neurotrophic activity of fibroblast growth factor-20 for dopaminergic neurons through fibroblast growth factor receptor-1c.* J Neurosci Res, 2003. **72**(4): p. 436-43.
90. Zhang, X., et al., *Serum FGF21 levels are increased in obesity and are independently associated with the metabolic syndrome in humans.* Diabetes, 2008. **57**(5): p. 1246-53.
91. Xu, J., et al., *Fibroblast growth factor 21 reverses hepatic steatosis, increases energy expenditure, and improves insulin sensitivity in diet-induced obese mice.* Diabetes, 2009. **58**(1): p. 250-9.
92. Inagaki, T., et al., *Inhibition of growth hormone signaling by the fasting-induced hormone FGF21.* Cell Metab, 2008. **8**(1): p. 77-83.
93. Palou, M., et al., *Sequential changes in the expression of genes involved in lipid metabolism in adipose tissue and liver in response to fasting.* Pflugers Arch, 2008. **456**(5): p. 825-36.
94. Arner, P., et al., *FGF21 attenuates lipolysis in human adipocytes - a possible link to improved insulin sensitivity.* FEBS Lett, 2008. **582**(12): p. 1725-30.
95. Umemori, H., et al., *FGF22 and its close relatives are presynaptic organizing molecules in the mammalian brain.* Cell, 2004. **118**(2): p. 257-70.
96. Sitara, D., et al., *Homozygous ablation of fibroblast growth factor-23 results in hyperphosphatemia and impaired skeletogenesis, and reverses hypophosphatemia in Phex-deficient mice.* Matrix Biol, 2004. **23**(7): p. 421-32.
97. Shimada, T., et al., *Targeted ablation of Fgf23 demonstrates an essential physiological role of FGF23 in phosphate and vitamin D metabolism.* J Clin Invest, 2004. **113**(4): p. 561-8.
98. Riminucci, M., et al., *FGF-23 in fibrous dysplasia of bone and its relationship to renal phosphate wasting.* J Clin Invest, 2003. **112**(5): p. 683-92.
99. Larsson, T., et al., *Transgenic mice expressing fibroblast growth factor 23 under the control of the alpha1(I) collagen promoter exhibit growth retardation, osteomalacia, and disturbed phosphate homeostasis.* Endocrinology, 2004. **145**(7): p. 3087-94.
100. Hesse, M., et al., *Ablation of vitamin D signaling rescues bone, mineral, and glucose homeostasis in Fgf-23 deficient mice.* Matrix Biol, 2007. **26**(2): p. 75-84.
101. Kuro-o, M., *Klotho as a regulator of fibroblast growth factor signaling and phosphate/calcium metabolism.* Curr Opin Nephrol Hypertens, 2006. **15**(4): p. 437-41.
102. Turnbull, J., A. Powell, and S. Guimond, *Heparan sulfate: decoding a dynamic multifunctional cell regulator.* Trends Cell Biol, 2001. **11**(2): p. 75-82.
103. Gallagher, J.T. and A. Walker, *Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and*

heparin are separate families of N-sulphated polysaccharides. Biochem J, 1985. **230**(3): p. 665-74.
104. Powell, A.K., et al., *Interactions of heparin/heparan sulfate with proteins: appraisal of structural factors and experimental approaches.* Glycobiology, 2004. **14**(4): p. 17R-30R.
105. McCormick, C., et al., *The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate.* Proc Natl Acad Sci U S A, 2000. **97**(2): p. 668-73.
106. Kobayashi, S., et al., *Association of EXT1 and EXT2, hereditary multiple exostoses gene products, in Golgi apparatus.* Biochem Biophys Res Commun, 2000. **268**(3): p. 860-7.
107. Hagner-Mcwhirter, A., U. Lindahl, and J. Li, *Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5.* Biochem J, 2000. **347 Pt 1**: p. 69-75.
108. Rosenberg, R.D., et al., *Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated?* J Clin Invest, 1997. **100**(11 Suppl): p. S67-75.
109. Kusche-Gullberg, M. and L. Kjellen, *Sulfotransferases in glycosaminoglycan biosynthesis.* Curr Opin Struct Biol, 2003. **13**(5): p. 605-11.
110. Rabenstein, D.L., *Heparin and heparan sulfate: structure and function.* Nat Prod Rep, 2002. **19**(3): p. 312-31.
111. Rudd, T.R. and E.A. Yates, *A highly efficient tree structure for the biosynthesis of heparan sulfate accounts for the commonly observed disaccharides and suggests a mechanism for domain synthesis.* Mol Biosyst, 2012. **8**(5): p. 1499-506.
112. Schlessinger, J., et al., *Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization.* Mol Cell, 2000. **6**(3): p. 743-50.
113. Ori, A., M.C. Wilkinson, and D.G. Fernig, *The heparanome and regulation of cell function: structures, functions and challenges.* Front Biosci, 2008. **13**: p. 4309-38.
114. Nunes, Q., *The role of Heparin-binding proteins in normal pancreas and acute pancreatitis*, 2015, University of Liverpool. p. 179.
115. Lin, X., *Functions of heparan sulfate proteoglycans in cell signaling during development.* Development, 2004. **131**(24): p. 6009-21.
116. Han, C., et al., *Distinct and collaborative roles of Drosophila EXT family proteins in morphogen signalling and gradient formation.* Development, 2004. **131**(7): p. 1563-75.
117. Yan, D. and X. Lin, *Shaping morphogen gradients by proteoglycans.* Cold Spring Harb Perspect Biol, 2009. **1**(3): p. a002493.
118. Sarrazin, S., W.C. Lamanna, and J.D. Esko, *Heparan sulfate proteoglycans.* Cold Spring Harb Perspect Biol, 2011. **3**(7).
119. Yayon, A., et al., *Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor.* Cell, 1991. **64**(4): p. 841-8.
120. Rapraeger, A.C., A. Krufka, and B.B. Olwin, *Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation.* Science, 1991. **252**(5013): p. 1705-8.
121. Zhu, H., et al., *The heparan sulfate co-receptor and the concentration of fibroblast growth factor-2 independently elicit different signalling patterns from the fibroblast growth factor receptor.* Cell Commun Signal, 2010. **8**: p. 14.

122. Lin, X., et al., *Heparan sulfate proteoglycans are essential for FGF receptor signaling during Drosophila embryonic development.* Development, 1999. **126**(17): p. 3715-23.
123. Delehedde, M., et al., *Fibroblast growth factor-2 binds to small heparin-derived oligosaccharides and stimulates a sustained phosphorylation of p42/44 mitogen-activated protein kinase and proliferation of rat mammary fibroblasts.* Biochem J, 2002. **366**(Pt 1): p. 235-44.
124. Faham, S., R.J. Linhardt, and D.C. Rees, *Diversity does make a difference: fibroblast growth factor-heparin interactions.* Curr Opin Struct Biol, 1998. **8**(5): p. 578-86.
125. Guimond, S., et al., *Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4.* J Biol Chem, 1993. **268**(32): p. 23906-14.
126. Fedarko, N.S. and H.E. Conrad, *A unique heparan sulfate in the nuclei of hepatocytes: structural changes with the growth state of the cells.* J Cell Biol, 1986. **102**(2): p. 587-99.
127. Ishihara, M., N.S. Fedarko, and H.E. Conrad, *Transport of Heparan-Sulfate into the Nuclei of Hepatocytes.* Journal of Biological Chemistry, 1986. **261**(29): p. 3575-3580.
128. Ishihara, M. and H.E. Conrad, *Correlations between Heparan-Sulfate Metabolism and Hepatoma Growth.* J Cell Physiol, 1989. **138**(3): p. 467-476.
129. Kovalszky, I., et al., *Inhibition of DNA topoisomerase I activity by heparan sulfate and modulation by basic fibroblast growth factor.* Molecular and Cellular Biochemistry, 1998. **183**(1-2): p. 11-23.
130. Sepulveda-Diaz, J.E., et al., *HS3ST2 expression is critical for the abnormal phosphorylation of tau in Alzheimer's disease-related tau pathology.* Brain, 2015. **138**: p. 1339-1354.
131. Raskind, W.H., et al., *Loss of Heterozygosity in Chondrosarcomas for Markers Linked to Hereditary Multiple Exostoses Loci on Chromosome-8 and Chromosome-11.* American Journal of Human Genetics, 1995. **56**(5): p. 1132-1139.
132. Hecht, J.T., et al., *Hereditary Multiple Exostosis and Chondrosarcoma - Linkage to Chromosome-11 and Loss of Heterozygosity for Ext-Linked Markers on Chromosome-11 and Chromosome-8.* American Journal of Human Genetics, 1995. **56**(5): p. 1125-1131.
133. Tornberg, J., et al., *Heparan sulfate 6-O-sulfotransferase 1, a gene involved in extracellular sugar modifications, is mutated in patients with idiopathic hypogonadotrophic hypogonadism.* Proc Natl Acad Sci U S A, 2011. **108**(28): p. 11524-11529.
134. Sleeman, M., et al., *Identification of a new fibroblast growth factor receptor, FGFR5.* Gene, 2001. **271**(2): p. 171-82.
135. Trueb, B., et al., *Characterization of FGFRL1, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues.* J Biol Chem, 2003. **278**(36): p. 33857-65.
136. Wang, F., et al., *Alternately spliced NH2-terminal immunoglobulin-like Loop I in the ectodomain of the fibroblast growth factor (FGF) receptor 1 lowers affinity for both heparin and FGF-1.* J Biol Chem, 1995. **270**(17): p. 10231-5.
137. Polanska, U.M., D.G. Fernig, and T. Kinnunen, *Extracellular interactome of the FGF receptor-ligand system: complexities and the relative simplicity of the worm.* Dev Dyn, 2009. **238**(2): p. 277-93.
138. Dailey, L., et al., *Mechanisms underlying differential responses to FGF signaling.* Cytokine & Growth Factor Reviews, 2005. **16**(2): p. 233-247.

139. Schlessinger, J., *Cell signaling by receptor tyrosine kinases.* Cell, 2000. **103**(2): p. 211-225.
140. Ornitz, D.M. and N. Itoh, *The Fibroblast Growth Factor signaling pathway.* Wiley Interdisciplinary Reviews-Developmental Biology, 2015. **4**(3): p. 215-266.
141. Yun, Y.R., et al., *Fibroblast growth factors: biology, function, and application for tissue regeneration.* J Tissue Eng, 2010. **2010**: p. 218142.
142. Teven, C.M., et al., *Fibroblast growth factor (FGF) signaling in development and skeletal diseases.* Genes Dis, 2014. **1**(2): p. 199-213.
143. Dorey, K. and E. Amaya, *FGF signalling: diverse roles during early vertebrate embryogenesis.* Development, 2010. **137**(22): p. 3731-42.
144. Spivak-Kroizman, T., et al., *Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation.* Cell, 1994. **79**(6): p. 1015-24.
145. Herr, A.B., et al., *Heparin-induced self-association of fibroblast growth factor-2. Evidence for two oligomerization processes.* J Biol Chem, 1997. **272**(26): p. 16382-9.
146. Moy, F.J., et al., *Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor.* Biochemistry, 1997. **36**(16): p. 4782-91.
147. Pellegrini, L., et al., *Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin.* Nature, 2000. **407**(6807): p. 1029-34.
148. Mach, H., et al., *Nature of the interaction of heparin with acidic fibroblast growth factor.* Biochemistry, 1993. **32**(20): p. 5480-9.
149. Sommer, A. and D.B. Rifkin, *Interaction of heparin with human basic fibroblast growth factor: protection of the angiogenic protein from proteolytic degradation by a glycosaminoglycan.* J Cell Physiol, 1989. **138**(1): p. 215-20.
150. Pantoliano, M.W., et al., *Multivalent ligand-receptor binding interactions in the fibroblast growth factor system produce a cooperative growth factor and heparin mechanism for receptor dimerization.* Biochemistry, 1994. **33**(34): p. 10229-48.
151. Springer, B.A., et al., *Identification and concerted function of two receptor binding surfaces on basic fibroblast growth factor required for mitogenesis.* J Biol Chem, 1994. **269**(43): p. 26879-84.
152. Sher, I., et al., *Mutations uncouple human fibroblast growth factor (FGF)-7 biological activity and receptor binding and support broad specificity in the secondary receptor binding site of FGFs.* J Biol Chem, 1999. **274**(49): p. 35016-22.
153. Harmer, N.J., et al., *Towards a resolution of the stoichiometry of the fibroblast growth factor (FGF)-FGF receptor-heparin complex.* J Mol Biol, 2004. **339**(4): p. 821-34.
154. Ibrahimi, O.A., et al., *Analysis of mutations in fibroblast growth factor (FGF) and a pathogenic mutation in FGF receptor (FGFR) provides direct evidence for the symmetric two-end model for FGFR dimerization.* Mol Cell Biol, 2005. **25**(2): p. 671-84.
155. Mohammadi, M., S.K. Olsen, and O.A. Ibrahimi, *Structural basis for fibroblast growth factor receptor activation.* Cytokine Growth Factor Rev, 2005. **16**(2): p. 107-37.
156. Ori, A., M.C. Wilkinson, and D.G. Fernig, *A systems biology approach for the investigation of the heparin/heparan sulfate interactome.* J Biol Chem, 2011. **286**(22): p. 19892-904.
157. Ishihara, M., et al., *Importance of 2-O-sulfate groups of uronate residues in heparin for activation of FGF-1 and FGF-2.* J Biochem, 1997. **121**(2): p. 345-9.

158. Rudland, P.S., et al., *Immunocytochemical identification of basic fibroblast growth factor in the developing rat mammary gland: variations in location are dependent on glandular structure and differentiation.* J Histochem Cytochem, 1993. **41**(6): p. 887-98.
159. Zhou, F.Y., et al., *Is the sensitivity of cells for FGF-1 and FGF-2 regulated by cell surface heparan sulfate proteoglycans?* Eur J Cell Biol, 1997. **73**(2): p. 166-74.
160. Turnbull, J.E., et al., *Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate.* J Biol Chem, 1992. **267**(15): p. 10337-41.
161. Habuchi, H., et al., *Structure of a heparan sulphate oligosaccharide that binds to basic fibroblast growth factor.* Biochem J, 1992. **285 (Pt 3)**: p. 805-13.
162. Kreuger, J., et al., *Interactions between heparan sulfate and proteins: the concept of specificity.* J Cell Biol, 2006. **174**(3): p. 323-7.
163. Thunberg, L., G. Backstrom, and U. Lindahl, *Further characterization of the antithrombin-binding sequence in heparin.* Carbohydr Res, 1982. **100**: p. 393-410.
164. Ornitz, D.M., et al., *Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells.* Mol Cell Biol, 1992. **12**(1): p. 240-7.
165. Yayon, A., et al., *Isolation of peptides that inhibit binding of basic fibroblast growth factor to its receptor from a random phage-epitope library.* Proc Natl Acad Sci U S A, 1993. **90**(22): p. 10643-7.
166. Walker, A., J.E. Turnbull, and J.T. Gallagher, *Specific heparan sulfate saccharides mediate the activity of basic fibroblast growth factor.* J Biol Chem, 1994. **269**(2): p. 931-5.
167. Ostrovsky, O., et al., *Differential effects of heparin saccharides on the formation of specific fibroblast growth factor (FGF) and FGF receptor complexes.* J Biol Chem, 2002. **277**(4): p. 2444-53.
168. Jemth, P., et al., *Biosynthetic oligosaccharide libraries for identification of protein-binding heparan sulfate motifs. Exploring the structural diversity by screening for fibroblast growth factor (FGF)1 and FGF2 binding.* J Biol Chem, 2002. **277**(34): p. 30567-73.
169. Xu, R., et al., *Analysis of the fibroblast growth factor receptor (FGFR) signalling network with heparin as coreceptor: evidence for the expansion of the core FGFR signalling network.* FEBS J, 2013. **280**(10): p. 2260-70.
170. Ori, A., et al., *Identification of heparin-binding sites in proteins by selective labeling.* Mol Cell Proteomics, 2009. **8**(10): p. 2256-65.
171. Sun, C., et al., *HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors.* PeerJ, 2015. **3**: p. e1060.
172. Kinsella, L., et al., *Interactions of putative heparin-binding domains of basic fibroblast growth factor and its receptor, FGFR-1, with heparin using synthetic peptides.* Glycoconj J, 1998. **15**(4): p. 419-22.
173. Seno, M., et al., *Carboxyl-terminal structure of basic fibroblast growth factor significantly contributes to its affinity for heparin.* Eur J Biochem, 1990. **188**(2): p. 239-45.
174. Rudd, T.R., et al., *Comparable stabilisation, structural changes and activities can be induced in FGF by a variety of HS and non-GAG analogues: implications for sequence-activity relationships.* Org Biomol Chem, 2010. **8**(23): p. 5390-7.
175. Uniewicz, K.A., et al., *Differential scanning fluorimetry measurement of protein stability changes upon binding to glycosaminoglycans: a screening test for binding specificity.* Anal Chem, 2010. **82**(9): p. 3796-802.

176. Rahmoune, H., et al., *Interaction of heparan sulfate from mammary cells with acidic fibroblast growth factor (FGF) and basic FGF. Regulation of the activity of basic FGF by high and low affinity binding sites in heparan sulfate.* J Biol Chem, 1998. **273**(13): p. 7303-10.
177. Delehedde, M., et al., *Fibroblast growth factor-2 stimulation of p42/44MAPK phosphorylation and IkappaB degradation is regulated by heparan sulfate/heparin in rat mammary fibroblasts.* J Biol Chem, 2000. **275**(43): p. 33905-10.
178. Duchesne, L., et al., *Transport of fibroblast growth factor 2 in the pericellular matrix is controlled by the spatial distribution of its binding sites in heparan sulfate.* PLoS Biol, 2012. **10**(7): p. e1001361.
179. Castellot, J.J., Jr., et al., *Binding and internalization of heparin by vascular smooth muscle cells.* J Cell Physiol, 1985. **124**(1): p. 13-20.
180. Uniewicz, K.A., et al., *Characterisation of the interaction of neuropilin-1 with heparin and a heparan sulfate mimetic library of heparin-derived sugars.* PeerJ, 2014. **2**: p. e461.
181. Migliorini, E., et al., *Cytokines and growth factors cross-link heparan sulfate.* Open Biol, 2015. **5**(8).
182. Wieth, J.O., P.J. Bjerrum, and C.L. Borders, Jr., *Irreversible inactivation of red cell chloride exchange with phenylglyoxal, and arginine-specific reagent.* J Gen Physiol, 1982. **79**(2): p. 283-312.
183. Zaki, L., *Anion transport in red blood cells and arginine-specific reagents. The location of [14C]phenylglyoxal binding sites in the anion transport protein in the membrane of human red cells.* FEBS Lett, 1984. **169**(2): p. 234-40.
184. Takahashi, K., *The reaction of phenylglyoxal with arginine residues in proteins.* J Biol Chem, 1968. **243**(23): p. 6171-9.
185. Konishi, K. and M. Fujioka, *Chemical modification of a functional arginine residue of rat liver glycine methyltransferase.* Biochemistry, 1987. **26**(25): p. 8496-502.
186. Ngo, T.T., et al., *p-Azidophenylglyoxal. A heterobifunctional photoactivable cross-linking reagent selective for arginyl residues.* J Biol Chem, 1981. **256**(21): p. 11313-8.
187. Politz, S.M., H.F. Noller, and P.D. McWhirter, *Ribonucleic acid-protein cross-linking in Escherichia coli ribosomes: (4-azidophenyl)glyoxal, a novel heterobifunctional reagent.* Biochemistry, 1981. **20**(2): p. 372-8.
188. Rostovtsev, V.V., et al., *A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes.* Angew Chem Int Ed Engl, 2002. **41**(14): p. 2596-9.
189. Greg, T.H., *Chapter 6 – Heterobifunctional Crosslinkers.* Bioconjugate Techniques (Third Edition), 2013.
190. Petitou, M., et al., *New synthetic heparin mimetics able to inhibit thrombin and factor Xa.* Bioorg Med Chem Lett, 1999. **9**(8): p. 1155-60.
191. Petitou, M., et al., *Synthesis of thrombin-inhibiting heparin mimetics without side effects.* Nature, 1999. **398**(6726): p. 417-422.
192. Pizette, S., et al., *Production and functional characterization of human recombinant FGF-6 protein.* Cell Growth Differ, 1991. **2**(11): p. 561-6.
193. Pizette, S. and L. Niswander, *BMPs negatively regulate structure and function of the limb apical ectodermal ridge.* Development, 1999. **126**(5): p. 883-94.
194. MacArthur, C.A., et al., *FGF-8 isoforms activate receptor splice forms that are expressed in mesenchymal regions of mouse development.* Development, 1995. **121**(11): p. 3603-13.
195. Jeffers, M., et al., *A novel human fibroblast growth factor treats experimental intestinal inflammation.* Gastroenterology, 2002. **123**(4): p. 1151-62.

196. Ron, D., et al., *Expression of biologically active recombinant keratinocyte growth factor. Structure/function analysis of amino-terminal truncation mutants.* J Biol Chem, 1993. **268**(4): p. 2984-8.
197. Gupta, A.A., et al., *Structural insights into the interaction of human S100B and basic fibroblast growth factor (FGF2): Effects on FGFR1 receptor signaling.* Biochim Biophys Acta, 2013. **1834**(12): p. 2606-19.
198. Rudd, T.R., et al., *Glycosaminoglycan origin and structure revealed by multivariate analysis of NMR and CD spectra.* Glycobiology, 2009. **19**(1): p. 52-67.
199. Rudd, T.R., et al., *Influence of substitution pattern and cation binding on conformation and activity in heparin derivatives.* Glycobiology, 2007. **17**(9): p. 983-93.
200. Bullock, S.L., et al., *Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase.* Genes Dev, 1998. **12**(12): p. 1894-906.
201. Merry, C.L., et al., *The molecular phenotype of heparan sulfate in the Hs2st-/- mutant mouse.* J Biol Chem, 2001. **276**(38): p. 35429-34.
202. Sun, C., Marcello, M., Li, Y., Mason, D., Lévy, R., Fernig, D. G.,, *Selectivity in glycosaminoglycan-binding dictates the distribution and diffusion of fibroblast growth factors in the pericellular matrix.* Open biology, 2015.
203. Sun, C. *Control of binding and movement of fibroblast growth factors by heparan sulfate in extracellular matrix*. [Thesis] 2015.
204. Jerabek-Willemsen, M., et al., *Molecular interaction studies using microscale thermophoresis.* Assay Drug Dev Technol, 2011. **9**(4): p. 342-53.