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

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

[1]Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK; [3]Laboratory of Cellular and Molecular Oncology, Centre de Recherche Public de la Santé (CRP-Santé), 84, Val Fleuri, L-1526 Luxembourg.

[2]Corresponding authors:

Changye Sun, Yong Li, David G. Fernig

Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

Tel. +44 151 795 4471

E. hscsun@liv.ac.uk, Y.li47@liv.ac.uk, dgfernig@liv.ac.uk

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. 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.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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 has been important in probing the mechanisms of, for example, their transport (Duchesne et al. 2012; Lin 2004; Yu et al. 2009). Chemical labeling has disadvantages compared to genetically encoded labelling, since with the latter it is easier to predict the structural and hence functional consequences of labeling, 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. 2006) 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 labeling.

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 et al. 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.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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); cDNAs 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. All 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, 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).

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 were cloned into pET-M11 such that the protein would have an N-terminal 6xhis tag followed by a and a tobacco etch virus (TEV) cleavage site (ENLYFQ) . FGF2 and FGF7 DNA sequences were previously cloned into pET-14b and pET-M11 (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 (Figure 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 (Figure 1: blue insert) was amplified by PCR using the Halo-FGF2-Forward, AAGGAGATATACCATGCCAGAAATCGGTACTG, and Halo-FGF2-Reverse, TCCCGGCTGCCATGGAGCTCTGAAAGTACAGATC, primers (NcoI and BamHI cleavage sites underlined), and inserted into the linearized vector using In-Fusion enzyme. A TEV cleavage site (Figure 1: green ellipsoid) was also included at the C-terminus of HaloTag to allow release of 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 (Figure 1).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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

His-FGF7, because it is toxic like native FGF7 (Sher et al. 2003), was transformed into BL21 (DE3) pLysS for subsequent protein expression and purification. FGF2, the other His-FGFs and Halo-FGFs were transformed into SoluBL21. 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 minutes 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 minutes. 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 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), all in 50 mM Tris-Cl buffer (pH 7.4). 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 chromatographies, were incubated with 2.5% (mol/mol) TEV protease at 4 °C overnight. In cases where the digestion products were cloudy, they were

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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 1 M NaCl in PB buffer and 0.1 M arginine, 1 M NaCl in PB buffer, respectively. After TEV digestion, FGF17 was further purified on a 1 mL HiTrap SP HP cation-exchange column by washing and eluting with 0.3 M NaCl and 1 M NaCl in 50 mM Tris-Cl buffer (pH 7.4). All of the fractions from the purification steps were analysed by 12% (w/v) SDS-PAGE.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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 (FGF1, FGF2 and FGF10), FGFs that expressed at a low level, FGF3 and FGF7, and FGFs that were insoluble when expressed in *E. coli* (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 2 A, C and E, 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, rather than in the pellet (Figs 2 A, C, E, lanes S and P). Chromatography of the supernatants on heparin demonstrated that little expressed protein was present in the flow-through fraction (Figs 2 A, C, E, lane T). Weak bands corresponding to His-FGF1 and His-FGF10, but not FGF2, were observed in the wash fraction (Figs 2 A, E, lane Wa), which may represent aggregated or less well-folded protein. The major proportion of the three FGFs was recovered in the high NaCl eluate (Figs 2 A, C and E, lane Hep), which indicated that these soluble FGFs bound heparin strongly and so 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 2 B, D and F, 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 2 B, D and F, lane 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 (Fig 2 B, D and F, 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 2 B, D and F, 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.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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 C, 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 3 A and C, lane Hep; Table 3).

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 3 B and D, lane L, red arrow) and in the soluble fraction obtained after centrifugation, whereas the pellet has relatively weaker bands (Figs 3 B and D, 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 3 B and D, 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 reported that FGF20 could also be solubilised by high concentrations of arginine (Maity et al. 2009), which suggests that FGF20 in the lysis buffer has a tendency to aggregate. Arginine would compete for binding of FGFs to heparin, however, 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 Buffer (Fig 5 A, lane 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-

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

FGF16 and Halo-FGF20 did not bind (Figs 5 B, C, D and E, lanes S and T). All of these four proteins could be recovered from heparin chromatography with the high concentration NaCl-containing elution buffers (Table 2) (Figs 5 B, C, D and E, lane Hep). When the Halo-FGF20 in the flow-through (Fig 5 F, 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 5 F, lane Hep2). A considerable amount of Halo-FGF16 also failed to bind to the heparin affinity column (Fig 5 E, lane T), though the bound protein was eluted with NaCl (Fig 5 E, 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, since nothing is known about the preference of either FGF16 or FGF20 for binding structures in the polysaccharide, if 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, the N-terminal HaloTag fusion clearly had a major effect on the expression of soluble protein. The yield of Halo-FGF6 and Hano-FGF20 was substantial (27 mg/L and 10 mg/L, respectively, Table 3). The lower yield of Halo-FGF8, Halo-FGF16, Halo-FGF17 and Halo-FGF22 (1 mg/L to 2 mg/L, Table 3) is 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, 5).

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. The eluates from heparin affinity chromatography of Halo-FGF1 and Halo-FGF7 were successfully purified by Q anion-exchange chromatography (Figs 6 A and B, 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 6 C and D, lane DEAE and CM). The isolated Halo-FGFs are quite pure, as is shown on the gels (Fig. 6).

Purification of FGFs by removing 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 7 A, red arrow) and the 18 kDa FGF2 (Fig 7 A, green arrow). Thus, the cleavage site is fully accessible to TEV protease. Both Halo-FGF17

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

and Halo-FGF20 were also well digested by TEV protease and subsequently soluble FGF17 (Fig 7 B, green arrow) and FGF20 (Fig 7 C, green arrow) were purified by cation-exchange and heparin chromatography, respectively.

Most of FGF6 (Fig 7 D, lane W_{Dig}, green arrow) and FGF22 (Fig 7 F, lane W_{Dig}, green arrow) and a small proportion of FGF8 were also released from HaloTag (Figs 7 D, E and F, 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 7 D, 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 7 D, lane E, green arrow). The disappearance of FGF8 and FGF22 in the soluble fractions after TEV digestion (Figs 7 E and F, lane S) showed that these two proteins were also not very soluble in the present buffer conditions without the HaloTag fusion partner.

Conclusion

In this study, we identified four useful properties of N-terminal HaloTag fusions for the production of 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 (Ferreira & Rasband 2012). 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. 2006) 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.

.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

313 **Acknowledgements**
314 Xianqing Mao wold like to thank Monika Dieterle for help with cloning Halo-FGF3.

Tables

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	~	MPEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHG NPTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYF FDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFHWAKRNP ERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRK LIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDRE PLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFW GTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPD LIGSEIARWLSTLEISGEPTTEDLYFQS-[FGFx]

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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	52.2	2.5	11
	No Tag		No Tag	
FGF3	28.2	60.0	0.5	11
FGF6	22.3	54.1	n.d.[1]	27
FGF7	22.2	54.0	0.6	5.6
FGF8	25.7	57.5	n.d.[1]	1.7
FGF10	22.7	54.5	7.7	9.3
FGF16	26.9	58.7	n.d.[1]	1.0
FGF17	25.8	57.6	n.d.[1]	1.5
FGF20	26.9	58.6	n.d.[1]	10
FGF22	20.5	52.3	n.d.[1]	2.0

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

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

Figure legends

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.

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-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; 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.

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.

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.

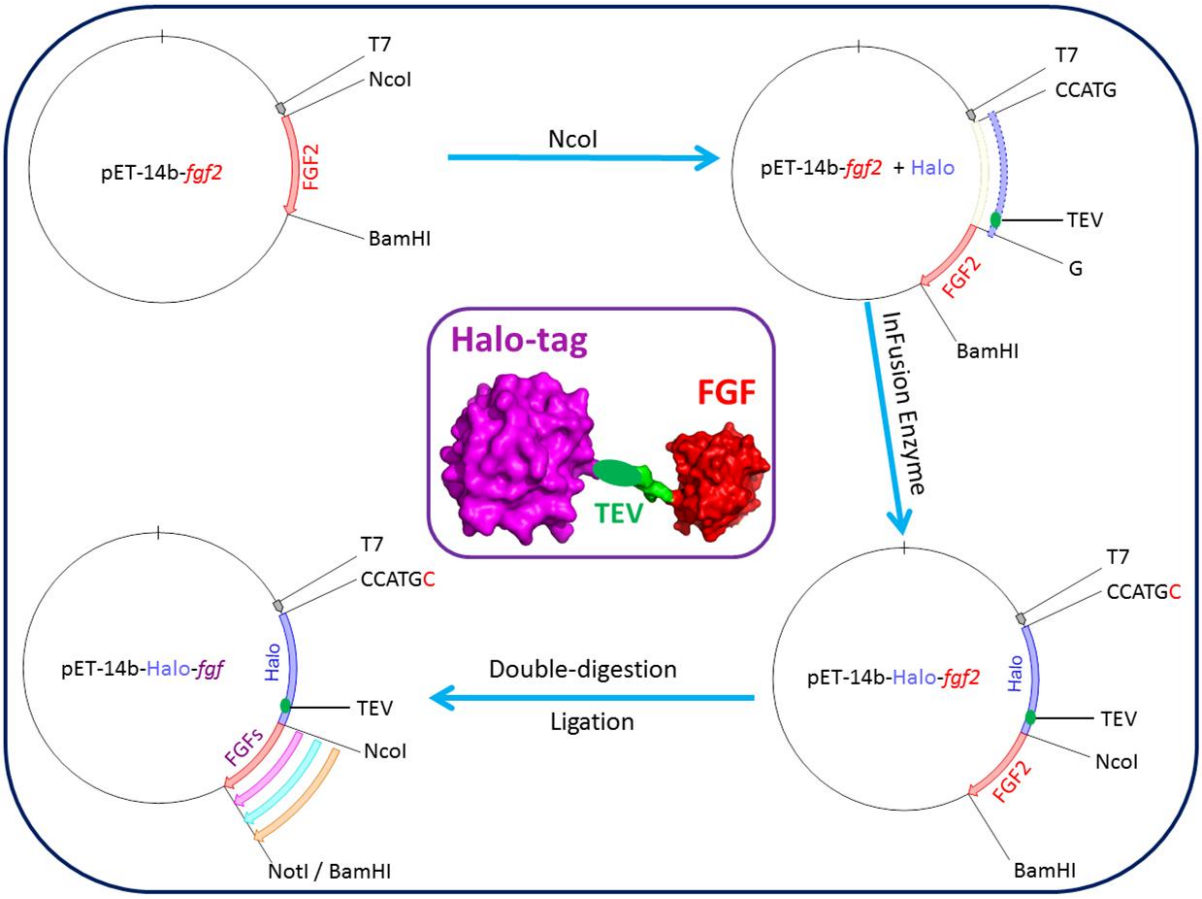
PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

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 as is shown above; 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 HaloTag and 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.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

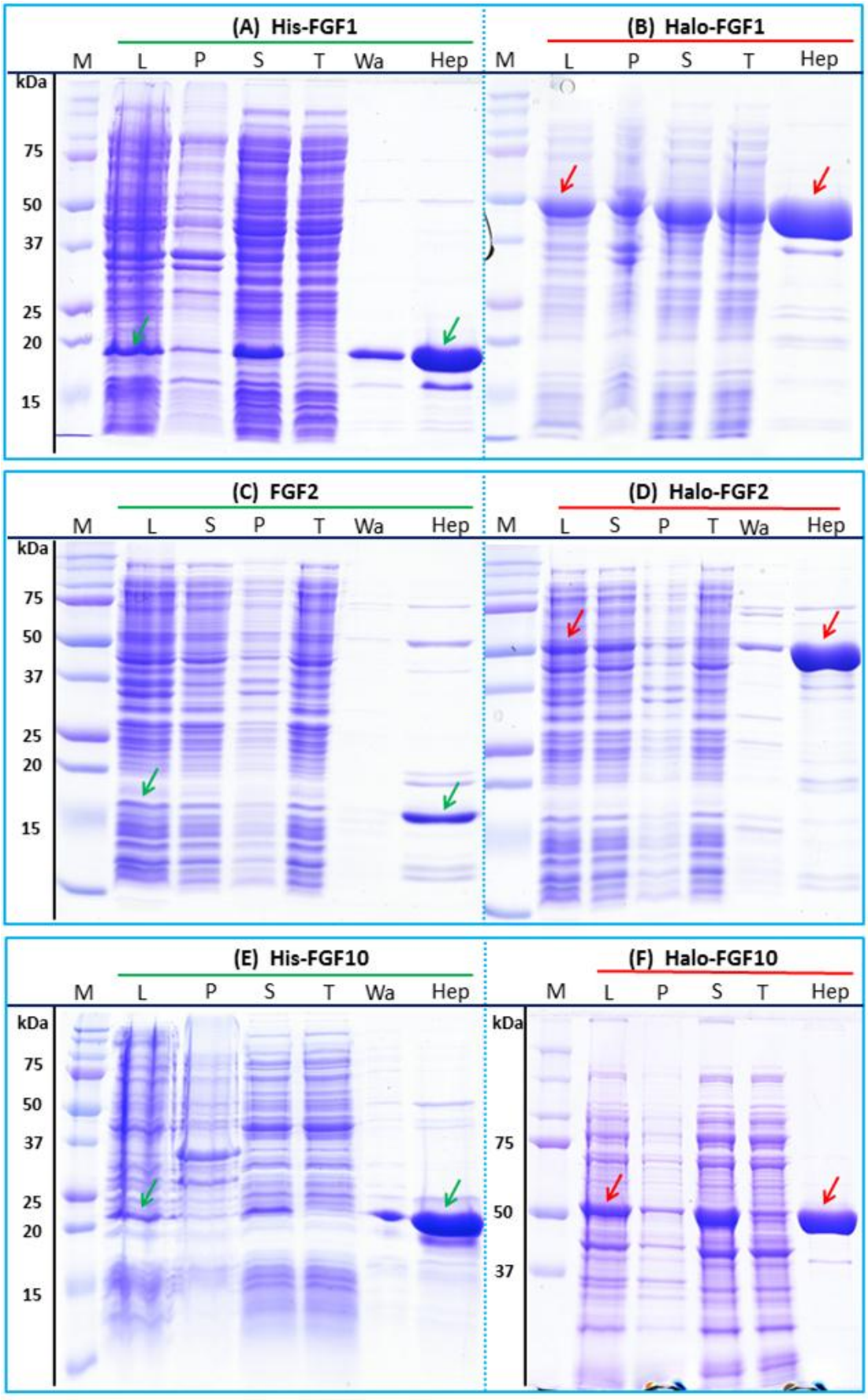
405 **Figure 1**



406

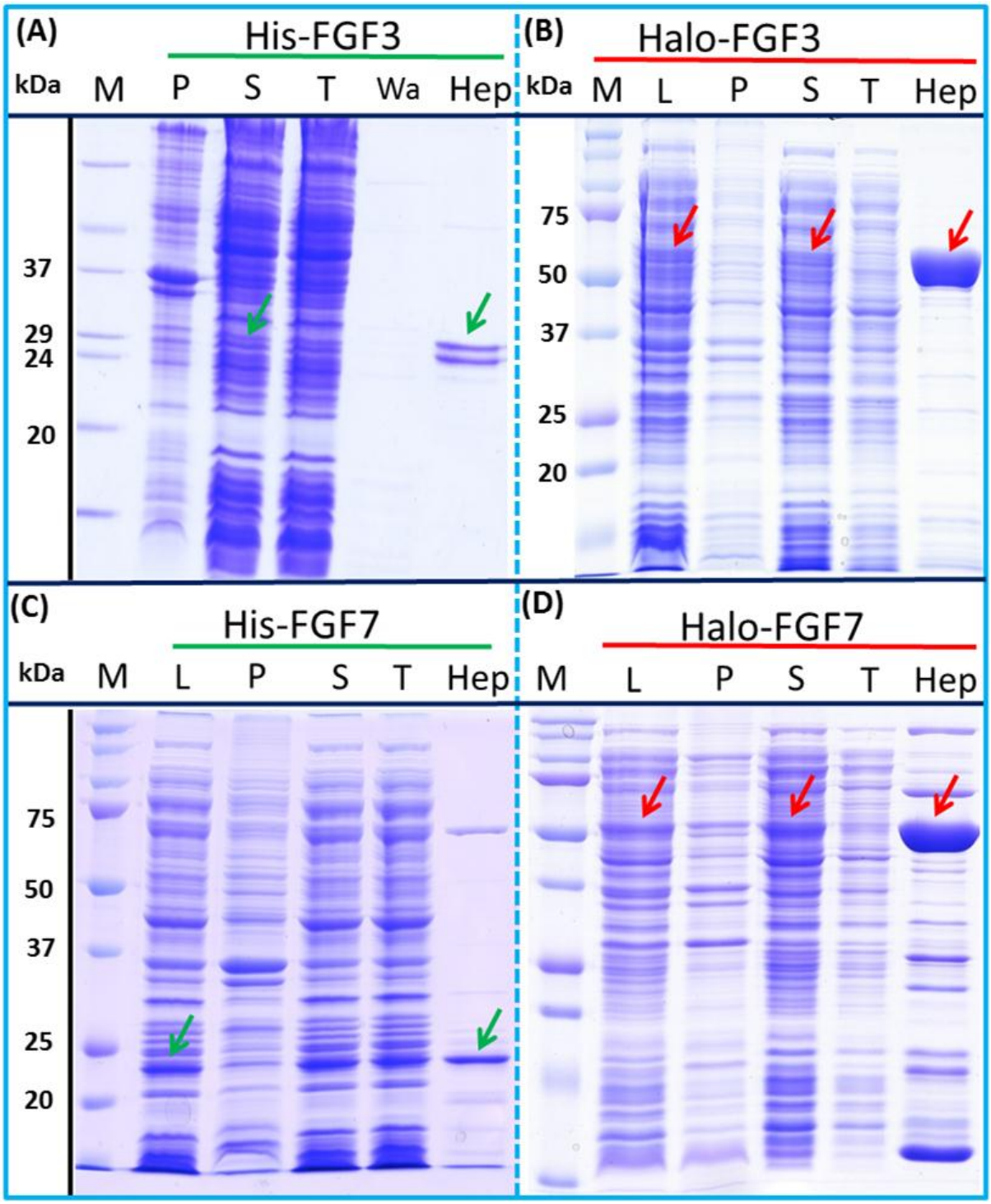
PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

Figure 2



PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

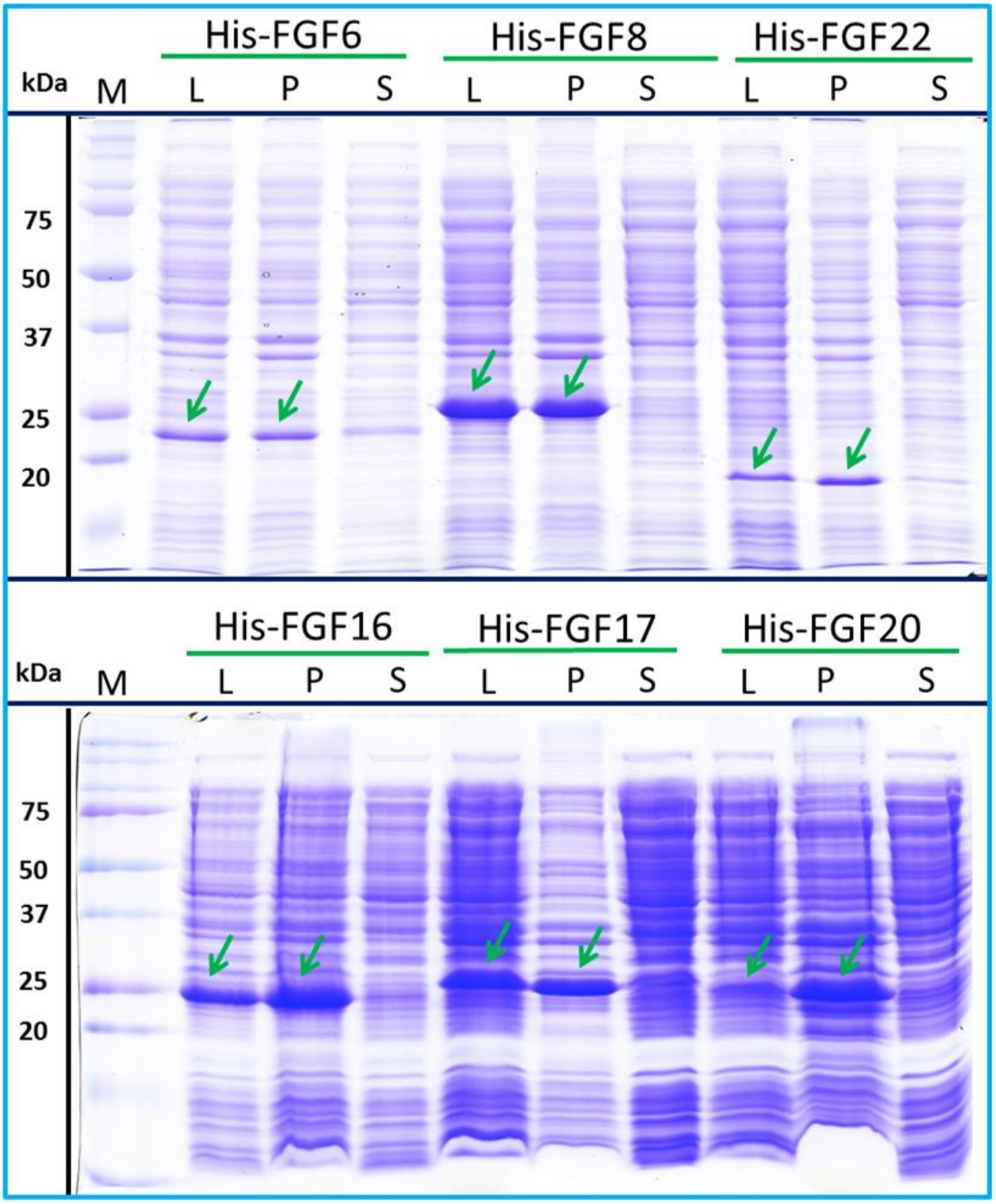
409 **Figure 3**



410
411

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

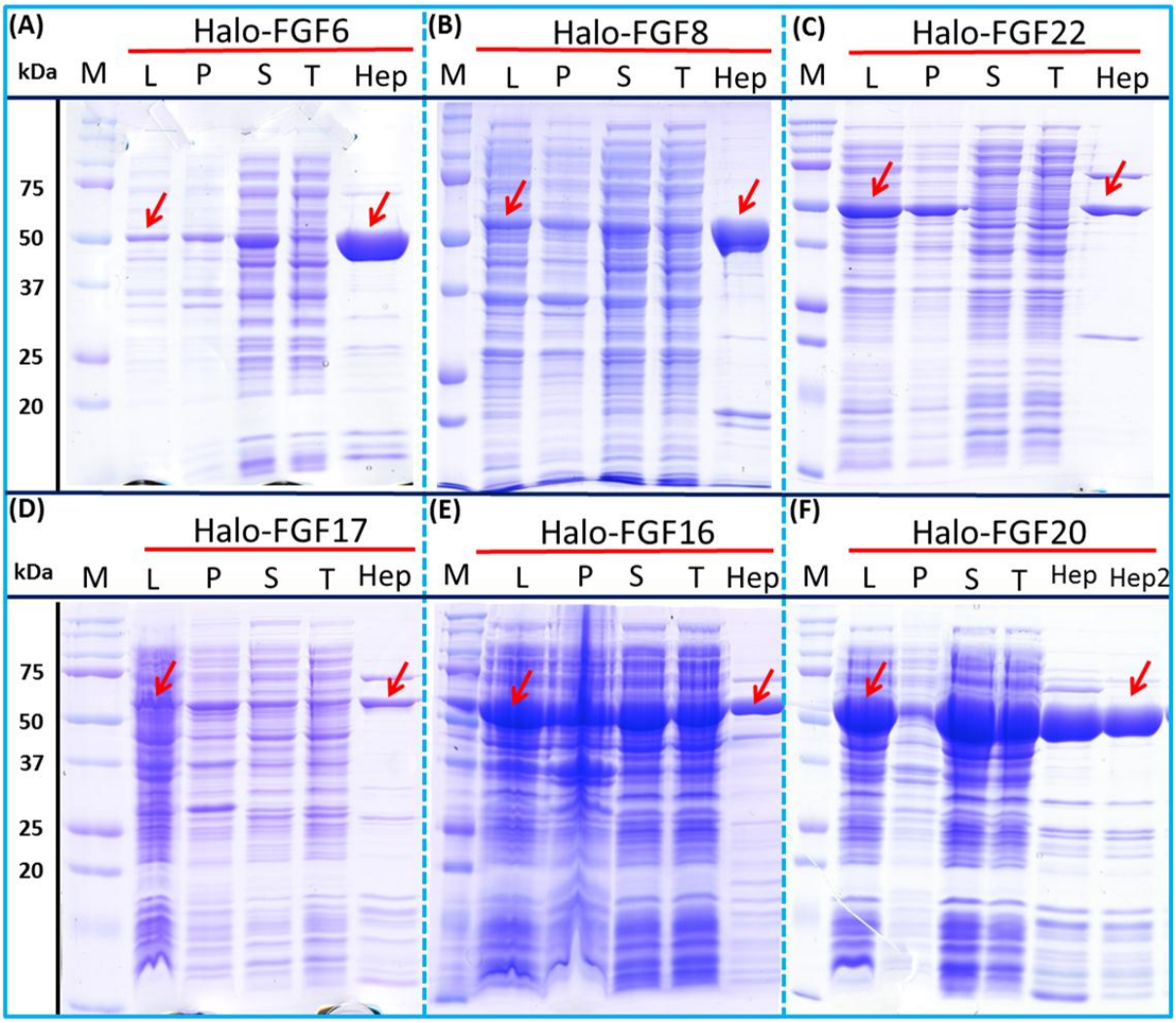
412 **Figure 4**



413
414

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

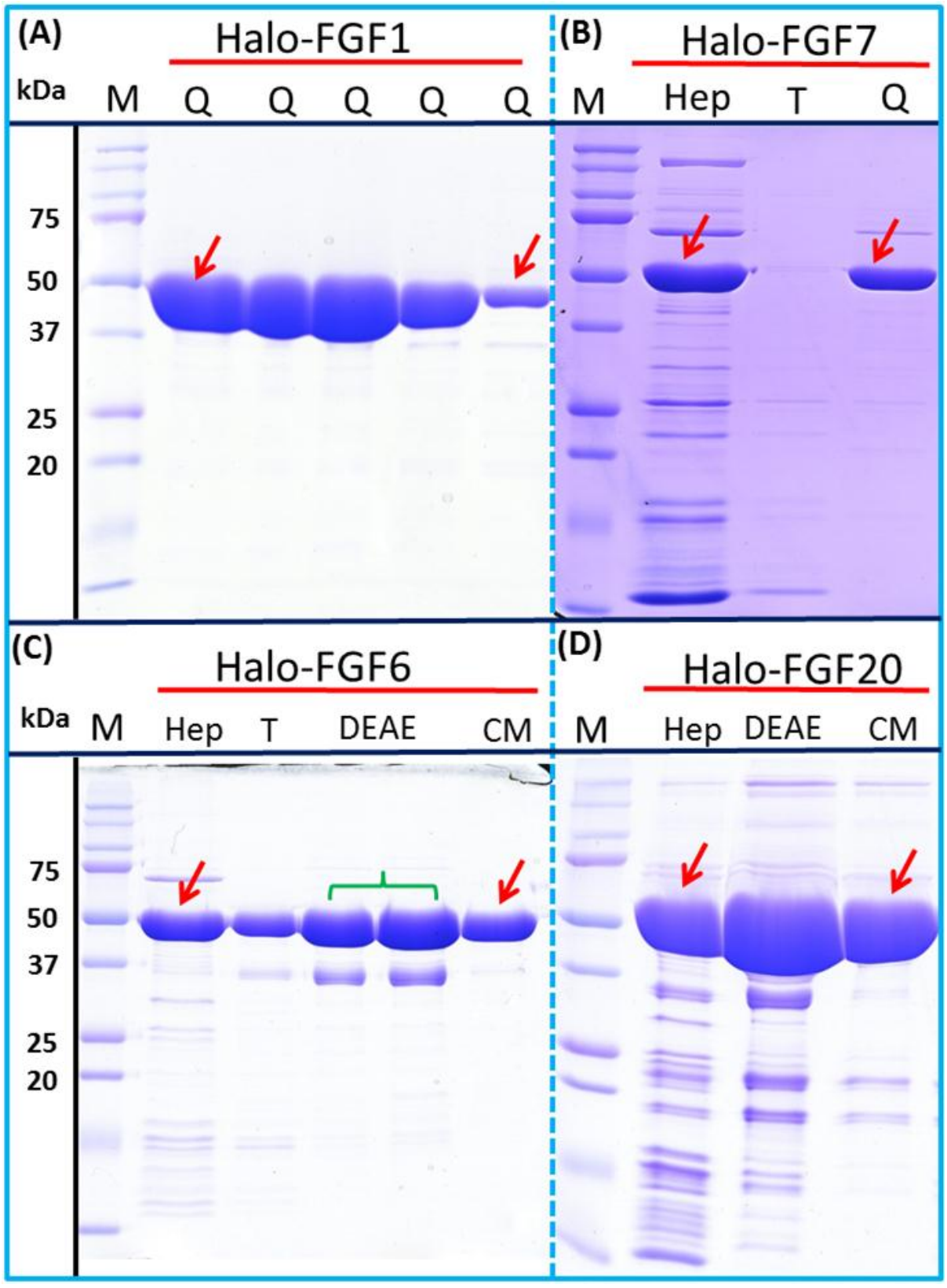
415 **Figure 5**



416

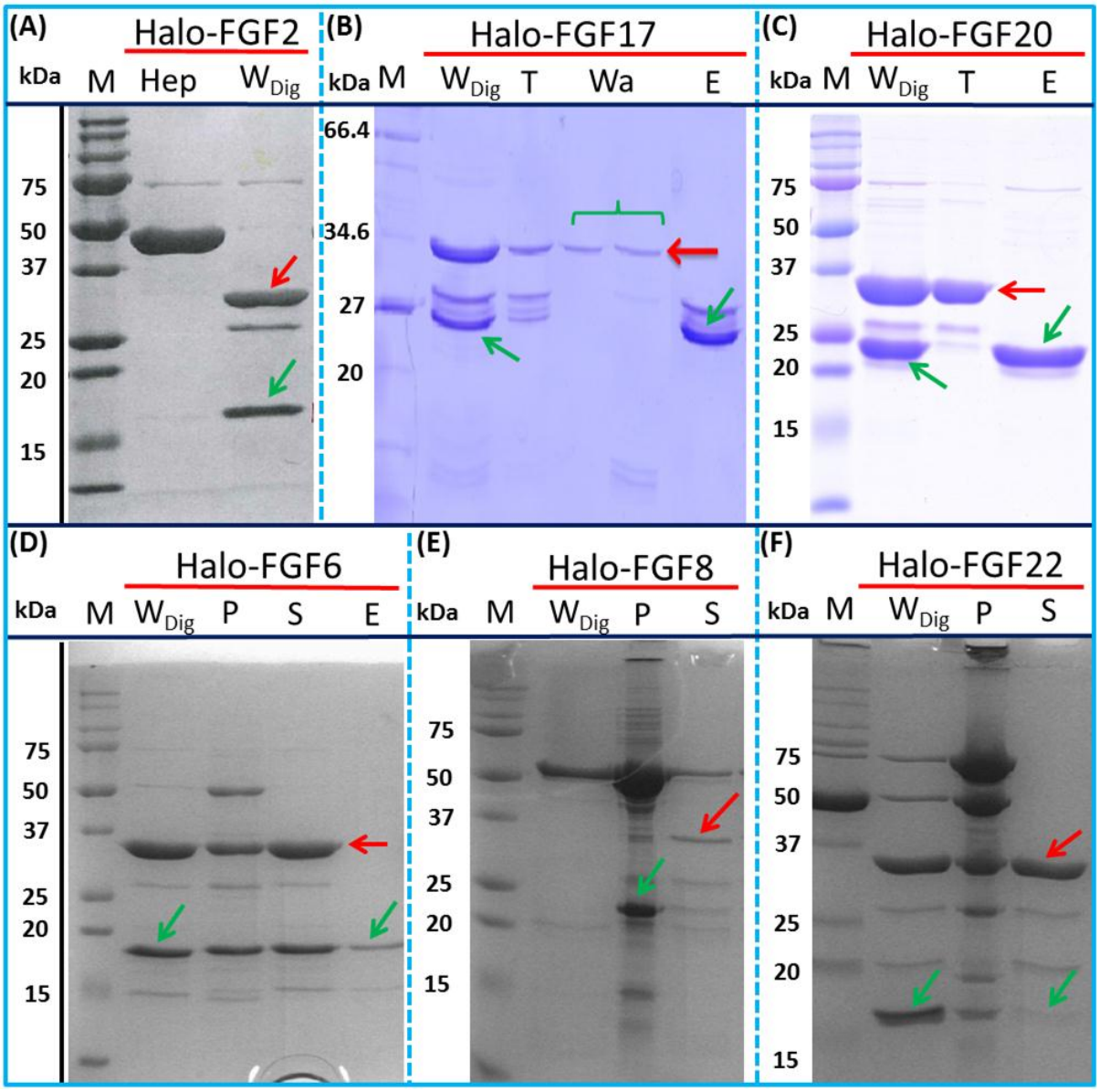
PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

Figure 6



PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

420 **Figure 7**



421

422

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

References

Asada M, Shinomiya M, Suzuki M, Honda E, Sugimoto R, Ikekita M, and Imamura T. 2009. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochimica Et Biophysica Acta-General Subjects* 1790:40-48.

Beenken A, and Mohammadi M. 2009. The FGF family: biology, pathophysiology and therapy. *Nature Reviews Drug Discovery* 8:235-253.

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, and Arakawa T. 1999. Recombinant rat fibroblast growth factor-16: Structure and biological activity. *Archives of Biochemistry and Biophysics* 361:34-46.

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

Duchesne L, Octeau V, Bearon RN, Beckett A, Prior IA, Lounis B, and 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.

Ferreira T, and 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, and Szoka FC. 2009. Tris-nitrilotriacetic acids of subnanomolar affinity toward hexahistidine tagged molecules. *Bioconjugate Chemistry* 20:1667-1672.

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.

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

Kalinina J, Byron SA, Makarenkova HP, Olsen SK, Eliseenkova AV, Larochelle WJ, Dhanabal M, Blais S, Ornitz DM, Day LA, Neubert TA, Pollock PM, and 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.

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

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

Lin XH. 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* 131:6009-6021.

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

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, and Wood KV. 2008. HatoTag: A novel protein labeling technology for cell imaging and protein analysis. *Acs Chemical Biology* 3:373-382.

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

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014

473 Maity H, Karkaria C, and Davagnino J. 2009. Effects of pH and arginine on the solubility and stability
474 of a therapeutic protein (fibroblast growth factor 20): relationship between solubility and
475 stability. *Current Pharmaceutical Biotechnology* 10:609-625.
476 Ohana RF, Encell LP, Zhao K, Simpson D, Slater MR, Urh M, and Wood KV. 2009. HaloTag7: A
477 genetically engineered tag that enhances bacterial expression of soluble proteins and
478 improves protein purification. *Protein Expression and Purification* 68:110-120.
479 Ornitz DM. 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development.
480 *Bioessays* 22:108-112.
481 Pizette S, Batoz M, Prats H, Birnbaum D, and Coulier F. 1991. Production and functional-
482 characterization of human recombinant FGF-6 protein. *Cell Growth and Differentiation*
483 2:561-566.
484 Ron D, Bottaro DP, Finch PW, Morris D, Rubin JS, and Aaronson SA. 1993. Expression of biologically-
485 active recombinant keratinocyte growth-factor - structure-function analysis of amino-
486 terminal truncation mutants. *Journal of Biological Chemistry* 268:2984-2988.
487 Roullier V, Clarke S, You C, Pinaud F, Gouzer G, Schaible D, Marchi-Artzner V, Piehler J, and Dahan M.
488 2009. High-affinity labeling and tracking of individual histidine-tagged proteins in live cells
489 using Ni^{2+} Tris-nitrilotriacetic acid quantum dot conjugates. *Nano Letters* 9:1228-1234.
490 Sher I, Yeg BK, Mohammadi M, Adir N, and Ron D. 2003. Structure-based mutational analyses in
491 FGF7 identify new residues involved in specific interaction with FGFR2IIIb. *Febs Letters*
492 552:150-154.
493 Susumu K, Medintz IL, Delehanty JB, Boeneman K, and Mattoussi H. 2010. Modification of
494 poly(ethylene glycol)-capped quantum dots with nickel nitrilotriacetic acid and self-assembly
495 with histidine-tagged proteins. *Journal of Physical Chemistry C* 114:13526-13531.
496 Tinazli A, Tang JL, Valiokas R, Picuric S, Lata S, Piehler J, Liedberg B, and Tampe R. 2005. High-affinity
497 chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: A
498 generic platform for protein chip technologies. *Chemistry-a European Journal* 11:5249-5259.
499 Turner N, and Grose R. 2010. Fibroblast growth factor signalling: from development to cancer.
500 *Nature Reviews Cancer* 10:116-129.
501 Uchinomiya S, Nonaka H, Fujishima S, Tsukiji S, Ojida A, and Hamachi I. 2009. Site-specific covalent
502 labeling of His-tag fused proteins with a reactive Ni(II)-NTA probe. *Chemical*
503 *Communications*:5880-5882.
504 Vogel A, Rodriguez C, and IzpisuaBelmonte JC. 1996. Involvement of FGF-8 in initiation, outgrowth
505 and patterning of the vertebrate limb. *Development* 122:1737-1750.
506 Xu RY, Ori A, Rudd TR, Uniewicz KA, Ahmed YA, Guimond SE, Skidmore MA, Siligardi G, Yates EA, and
507 Fernig DG. 2012. Diversification of the structural determinants of fibroblast growth factor-
508 heparin interactions implications for binding specificity. *Journal of Biological Chemistry*
509 287:40061–40073.
510 Yu S, Burkhardt M, Nowak M, Ries J, Petrásek Z, Scholpp S, Schwille P, and Brand M. 2009. FGF8
511 morphogen gradient is formed by a source-sink mechanism with freely diffusing molecules.
512 *Nature* 461:533-536.
513 Zhang Y, So MK, Loening AM, Yao HQ, Gambhir SS, and Rao JH. 2006. HaloTag protein-mediated site-
514 specific conjugation of bioluminescent proteins to quantum dots. *Angewandte Chemie-*
515 *International Edition* 45:4936-4940.

516

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.743v1 | CC-BY 4.0 Open Access | rec: 20 Dec 2014, publ: 20 Dec 2014