Sequence analyses and comparative modeling of fly and worm fibroblast growth factor receptors indicate that the determinants for FGF and heparin binding are retained in evolution

H.G. Nagendra*, Adrian E. Harrington, Nicholas J. Harmer, Luca Pellegrini, Tom L. Blundell[1], David F. Burke

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

Abstract The presence of a large number of fibroblast growth factors (FGFs) and multiple splice forms of their receptors (FGFRs) in higher vertebrates makes the three-dimensional (3D) analysis of FGF interactions with their receptors a formidable task. The situation differs in *Caenorhabditis elegans* (worm) and *Drosophila melanogaster* (fruit fly), where only one or two FGF and FGFR sequences have been identified. Structural studies of the FGF–FGFR complexes in such primitive organisms should reveal the basic features of the ligand–receptor interactions as they first emerged through evolution. We have analysed the sequences of worm and fly FGFs and FGFRs and used the recently determined crystal structure of the human FGF1–FGFR2–heparin ternary complex [Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B. and Blundell, T.L. (2000) Nature 407, 1029–34] to construct 3D models of the homologous complexes. In spite of a low sequence similarity with their human counterparts, key structural features required for ligand–receptor and protein–heparin binding in humans are conserved in the fly and worm FGF–FGFR–heparin complexes. Analyses of the models show that tertiary interactions that are not conserved in sequence are maintained through novel interactions or complementary mutations in the fly and worm sequences. The overall charge distributions observed in the human FGF–FGFR–heparin complex are retained in the fly and worm models. The arginine residue at position 253 in the linker region between the Ig-like domains D2 and D3 in the wild type fly and worm sequences is particularly striking, as the Pro253Arg mutation in humans is responsible for Apert syndrome. This change may enhance the affinity of receptors for their FGF molecules as observed in Apert mutants. © 2001 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Fibroblast growth factor receptor; Heparin binding; Apert syndrome; *Caenorhabditis elegans*; *Drosophila melanogaster*

1. Introduction

Fibroblast growth factors (FGFs) comprise a large family of growth factors that are key regulators of cell proliferation, migration and differentiation [2–4], with vital roles not only in normal development and wound healing [5–9], but also in tumour development and progression [10]. FGF receptors (FGFRs) consist of up to three extracellular immunoglobulin (Ig)-like domains (D1, D2 and D3), a single transmembrane chain and an intracellular domain with tyrosine kinase activity [11]. Upon FGF binding and receptor dimerisation, the kinase domains of FGFRs undergo cross-phosphorylation, resulting in signal transduction and biological response [12–14]. The integrity of these developmental pathways requires extremely tight regulation of FGF activity [15–19], and the essential presence of heparin or heparan sulphate for the formation of an active FGF signaling complex [20–32].

Twenty-two FGFs and several FGFR splice variants, arising from alternative mRNA splicing of four genes, have been found in humans, and a similar variety of FGFs and their receptors is widespread in the animal kingdom. However, in the *Caenorhabditis elegans* genome, there are only two identified FGF homologues (egl17 and let756) and one FGFR sequence (egl15). Let756 was detected from the sequence data [33], while egl15 and egl17 were characterised by genetic and biochemical screens on the egg-laying defective (egl) mutants, implicating them in sex myoblast migration during development [34,35]. In the *Drosophila melanogaster* genome, one FGF homologue [15] and two FGFR proteins, DFR1 and DFR2 [36], have been described.

The recently published crystal structures of the human ternary complexes FGF1–FGFR2–heparin [1] and FGF2–FGFR1–heparin [37] exhibit both common and unique features. The former (PDB code: 1e0o) is a co-crystallised 2:2:1 ternary complex, and the latter structure (PDB code: 1fq9) a 2:2:2 complex resulting from soaking of heparin fragments in FGF2–FGFR1 crystals. Interestingly, the two structures share common heparin–protein interactions (Fig. 1). However, in 1e0o, heparin bridges the two 1:1 FGF:FGFR complexes, but in 1fq9 it is directly involved in receptor–receptor dimerisation. The differences in the relative orientation of the domain D3 of the receptor may be induced by heparin binding [1].

The structure of the human FGF–FGFR–heparin complex allows us to consider whether the homologues in *D. melanogaster* and *C. elegans* can form similar ternary complexes. The comparative modeling studies on the FGF–FGFR–heparin complexes described here indicate a conserved heparin-binding pocket, consistent with the observation that heparan

*Corresponding author. Fax: (44)-1223-766082.
E-mail: tom@cryst.bioc.cam.ac.uk

[1] On leave from Sir M.V. Institute of Technology, Bangalore 562 157, India.

Fig. 1. Structural alignment of various models generated by COMPARER [48] indicating the heparin-binding residues (shaded in yellow): (A) FGFs and (B) FGFRs.

sulphate has a direct role in FGFR activation in *D. melano-gaster* development [20]. The models suggest that the key determinants for FGF and heparin binding are retained in the fly and worm sequences, although specific amino acid interactions observed in the human structure may not be conserved between the FGFRs and their ligands.

2. Materials and methods

The alignment of the FGF and FGFR sequences from *D. mela-nogaster* and *C. elegans* with the human sequences present in the structure of Pellegrini et al. [1], was performed using the programs FUGUE [38], CLUSTALW [39], SEAVIEW [40] and JOY [41]. These alignments were used as input to Modeller4 [42,43] to generate initial three-dimensional (3D) models of the FGF–FGFR–heparin complexes, with the heparin coordinates from 1e0o. The template structure 1e0o has 15 disordered residues in each of its receptor chains and the corresponding regions in the fly and worm sequences were omitted during modeling. The program CODA [44] was used to remodel the loop conformations and the side chains for these remodeled loops were built by SCWRL [45]. The graphical software packages 'O' [46] and Swisspdb Viewer [47] were employed to check appropriate rotamer conformations of the side chain atoms, in order to optimise possible interactions. The models were then subjected to another round of energy minimisation using Modeller4 and further assessed by examining the JOY formatted sequences (Fig. 1), aligned on the basis of their structures using the program COMPARER [48]. The final models were also checked for geometrical acceptability with the programs PROCHECK [49] and VERIFY3D [50]. The program CONTACT in the CCP4 package [51] was used to identify atoms involved in protein–heparin and protein–protein contacts. The cut-off distance for hydrogen bonding or hydrophobic contacts was set to 4.0 Å, to accommodate all plausible interactions.

3. Results and discussion

3.1. Analysis of sequences and 3D models

Although the sequence identities between the human, fly and worm sequences are in the range 22–35% (Table 1), many of the residues previously reported as completely conserved within the FGF and FGFR family [52] are also present in the fly and worm sequences. This allowed models of the fly and worm complexes to be constructed with good geometry, and PROCHECK showed that more than 90% of residues for each model were within the allowed regions of their ϕ–ψ space [49]. The values of the root mean square deviations (RMSD) of the C-α superposition of the fly and worm models with 1e0o (Table 1), indicate that the core structures in the models are well conserved. In Fig. 1, the residue numbering refers to the models as FGF–DFR1–heparin (Fly1), FGF–DFR2–heparin (Fly2), Egl17–egl15–heparin (Worm1), let756–egl15–heparin (Worm2) and the template human ternary complex (1e0o). However, in all ensuing discussions we follow the residue numbering used in the template X-ray structure (1e0o),

Table 1
Percentage identities among the sequences and the RMSD between their α-carbon positions in various models

RMSDs	PIDs				
	1e0o	Fly1	Fly2	Worm1	Worm2
1e0o	–	32.9	29.7	27.3	35.0
Fly1	0.87	–	62.1	23.5	31.1
Fly2	0.99	0.96	–	22.1	29.8
Worm1	0.88	0.93	1.20	–	68.4
Worm2	0.43	0.98	1.05	0.97	–

Table 2
Residues in various models involved in equivalent interactions across ligand–receptor interfaces (residues marked with * indicate a possibility of an interaction equivalent to that seen in 1e0o)

1e0o	Fly1	Fly2	Worm1	Worm2
Molecule A				
Y15	Y6	Y6	F6	F6
R37	E26	E26	K29	E36
V51	V40	V40	D42	M50
L89	M78	M78	V81	M90
Y94	Y83*	Y83*	F86	Y95
L133	Y125	Y125	L125	L136
P134	T126	T126	V126	V137
L135	N127	N127	R127	V138*
Molecule B				
L166	L147	L141	T150	T161
A168	T149	H143	A152	A163
P170	P151	L145	P154	P165
D247	I225	Q216	I232	I243
R251	R229	R220	R236	R247
I257	I235	I226	I242	I253
I350	Q315	R312	M327	M338
Molecule C				
Y15	Y331	Y328	F344	F355
S17	K333	K330	E346*	R357
R35	Q349	Q346	P365*	K383
L89	M403	M400	V419	M439
E90	G404	G401	R420	M440
N92	Q406	Q403	N422	N442
H93	N407	N404	F424	Y443
L133	Y450	Y447	L463	L485
Molecule D				
R165	F471	R462	E487	E509
L166	L472	L463	T488	T510
A168	T474	H465	A490	A512
V169	R475	S466	L491	L513
P170	P476	L467	P492	P514
R251	R554	R542	R574	R596
R255	K538	A546	P578	P600
I257	I560	I548	I580	I602

Molecules A and C = FGFs; Molecules B and D = FGFRs.

i.e. ligand: 10–138; FGFR-D2: 149–250; linker region: 251–254; FGFR-D3: 255–360.

Alignment with all currently available FGF sequences showed that the residues that are absolutely conserved [52] are retained in let756 and the fly FGF. However in egl17 the invariant M67 is replaced by phenylalanine and G71 by alanine. The substitution of a large hydrophobic side chain in place of methionine is easily accommodated in the core of the ligand. Other features of egl17 include the presence of five tryptophan residues and a higher percentage of basic residues in the core region.

3.2. Receptor–ligand interactions

Analyses of the receptor–ligand contacts in Fly1, Fly2, Worm1 and Worm2 complexes suggest that they are likely to be similar to those of human homologues [1,53,54]. A total of 32 residues are involved in different types of interactions across the interfaces of ligand–receptor complexes in 1e0o, and between 13 and 18 of these interactions are conserved between the models and the template structure 1e0o. The residues involved in equivalent ligand–receptor contacts in the models are listed in Table 2.

Interactions between the FGF and FGFR domain D2 are primarily hydrophobic, involving the residues aligned with Y15, Y94, L133, P134, L135 in the ligand and L166, A168 and P170 in the receptor. Contacts between the ligand and

domain D3 are also hydrophobic with a well conserved interaction between V51 and I350. Involvement of R255 (in D3) in an electrostatic interaction with an invariant glutamate (E87) in the ligand has been demonstrated to be important, as a 10-fold reduction in receptor binding affinity is observed when this glutamate is mutated to glutamine [55]. The absence of this charge–charge interaction in the Fly2 and worm models, chiefly due to substitutions in position 255, could suggest an alternate mode of binding and a different orientation of domain D3. However, an interaction with a substituted lysine at position R255 in Fly1 appears feasible.

There exist some differences with potential functional consequences in the residues linking the domains D2 and D3 (Fig. 1). Two point mutations in FGFR2, S252W or P253R, have been shown to be responsible for the vast majority of cases of Apert syndrome, a human congenital condition characterised by craniosynostosis and severe syndactly of the fingers and toes [13,56]. In vitro studies [57,58] on mutant receptors have shown increased ligand–receptor affinity, a broadening of receptor specificity, and an increased lifespan of the ligand–receptor complex. Intriguingly, the fly and worm receptors have an arginine at the position equivalent to P253, a feature of Apert [13], Pfeiffer [59] and Muenke [60] syndromes, which could have implications for the affinity of the receptor for ligand in the fly and worm complexes and may provide better stabilising interactions with the FGF ligand [57,58].

The network of interactions of the invariant R251 in the linker region gains significance as mutagenesis data on a ligand residue (N95A) reveal a 400-fold reduction in receptor affinity [61]. In the human complex 1e0o, the guanidinium moiety of invariant R251 in the linker between domains D2 and D3 is surrounded by three hydrophobic ligand residues L89, L133 and P134, and is hydrogen bonded to the main chain oxygen of H93 and the side chain oxygen of N95 [1]. The sequence alignment (Fig. 1) indicates the conservation of this asparagine in fly and let756 ligands, while it is substituted to a serine in egl17. However, a similar network of interactions is retained in both the fly and worm models. In Fly1 and Fly2, the invariant arginine residue equivalent to R251 makes a strong hydrogen bond with a threonine residue, present in the position equivalent to P134, while the neighbouring asparagines, equivalent to H93 and N95, appear to maintain the interaction geometry. The methionine and tyrosine residues aligned at positions L89 and P133 may well contribute to equivalent hydrophobic interactions. Likewise in the worm models, the invariant residue at R251 makes very similar hydrogen bonds with ligand residues in positions equivalent to H93 and N95. The hydrophobic interactions rendered by L89, L133, and P134 appear well maintained by the equivalent residues in the worm ligands. Thus, the common features between ligand–receptor interactions observed in fly and worm models and those in the human receptor complex suggest a similar mode of recognition.

3.3. Protein–heparin interactions

The X-ray structure of the human FGF1–FGFR2–heparin ternary complex [1] reveals that the heparin decasaccharide mediates the dimerisation of two 1:1 ligand–receptor complexes, lying between and interacting with both FGF molecules and one receptor molecule. A range of basic and polar residues in FGF and FGFR constitute the binding sites for heparin (Table 3). The analysis of equivalent heparin-binding
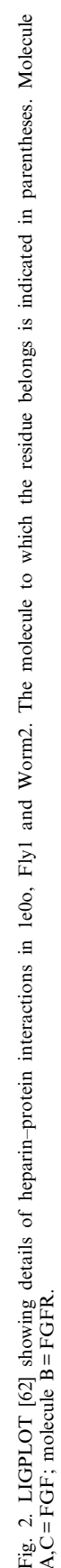
Table 3
Residues in equivalent positions interacting with the heparin molecule in various models

1e0o	Fly1	Fly2	Worm1	Worm2
Molecule A				
–	K8	K8	–	–
N18	N9	N19	Q11	S9
–	–	Y99	–	–
–	L100	–	–	–
K112	A101	A101	N102	R115
K113	L102	L102	G103	R116
N114	N103	N103	R104	S117
–	–	–	R106	–
–	Q107	Q107	–	–
–	P108	–	–	–
K118	R109	R109	Q108	R121
R119	R110	R110	N109	R122
–	–	–	Y113	–
R122	I113	I113	H114	N125
G126	R117	R117	R118	R129
Q127	S118	S118	C119	R130
K128	L119	L119	F120	K131
A129	G120	G120	D121	A132
Molecule B				
–	K142	–	–	–
K161	K143	–	N143	N154
–	–	E137	–	–
K164	–	K139	–	–
H167	Q148	Q142	H151	H162
–	–	N148	R157	–
T174	L155	T149	T158	T169
V175	L156	V150	L159	L170
K176	T157	N151	K160	K171
–	–	–	–	R175
R178	–	–	N162	–
–	K188	–	–	–
–	R191	–	–	–
I217	–	R186	–	–
–	–	–	–	E213
Molecule C				
–	–	–	R399	–
–	–	–	R401	K420
K105	A419	A416	–	R455
–	R420	R417	–	–
W107	R421	R418	–	W459
K112	–	–	–	R464
–	–	–	R442	–
–	–	–	–	K468
–	–	Q429	–	–
–	P433	–	–	–
–	–	R431	–	R471
R119	R435	R432	L449	–
–	–	–	S450	–
P121	Q437	Q434	Y451	P473
R122	I438	I435	H452	N474

Molecules A and C = FGFs; Molecule B = FGFR (heparin does not interact with Molecule D).

residues in the *D. melanogaster* and *C. elegans* sequences indicates high similarity. Residues that are involved in interaction with the heparin molecule in the models of 1e0o, Fly1, Fly2, Worm1 and Worm2 respectively are highlighted in Fig. 1 and listed in Table 3.

A total of 25 residues in 1e0o interact with the heparin molecule. In both models of Fly1 and Fly2, 30 residues were identified as interacting with heparin. Of these, 16 residues in 1e0o, 22 in Fly1 and 24 in Fly2 interact by forming hydrogen bonds, the majority of which (14 in 1e0o, 11 in Fly1 and 10 in Fly2) are basic residues. The remaining residues are involved in hydrophobic contacts. Likewise, 27 and 26 residues interact with the heparin molecule in Worm1 and

Fig. 2. LIGPLOT [62] showing details of heparin–protein interactions in 1e0o, Fly1 and Worm2. The molecule to which the residue belongs is indicated in parentheses. Molecule A,C = FGF; molecule B = FGFR.
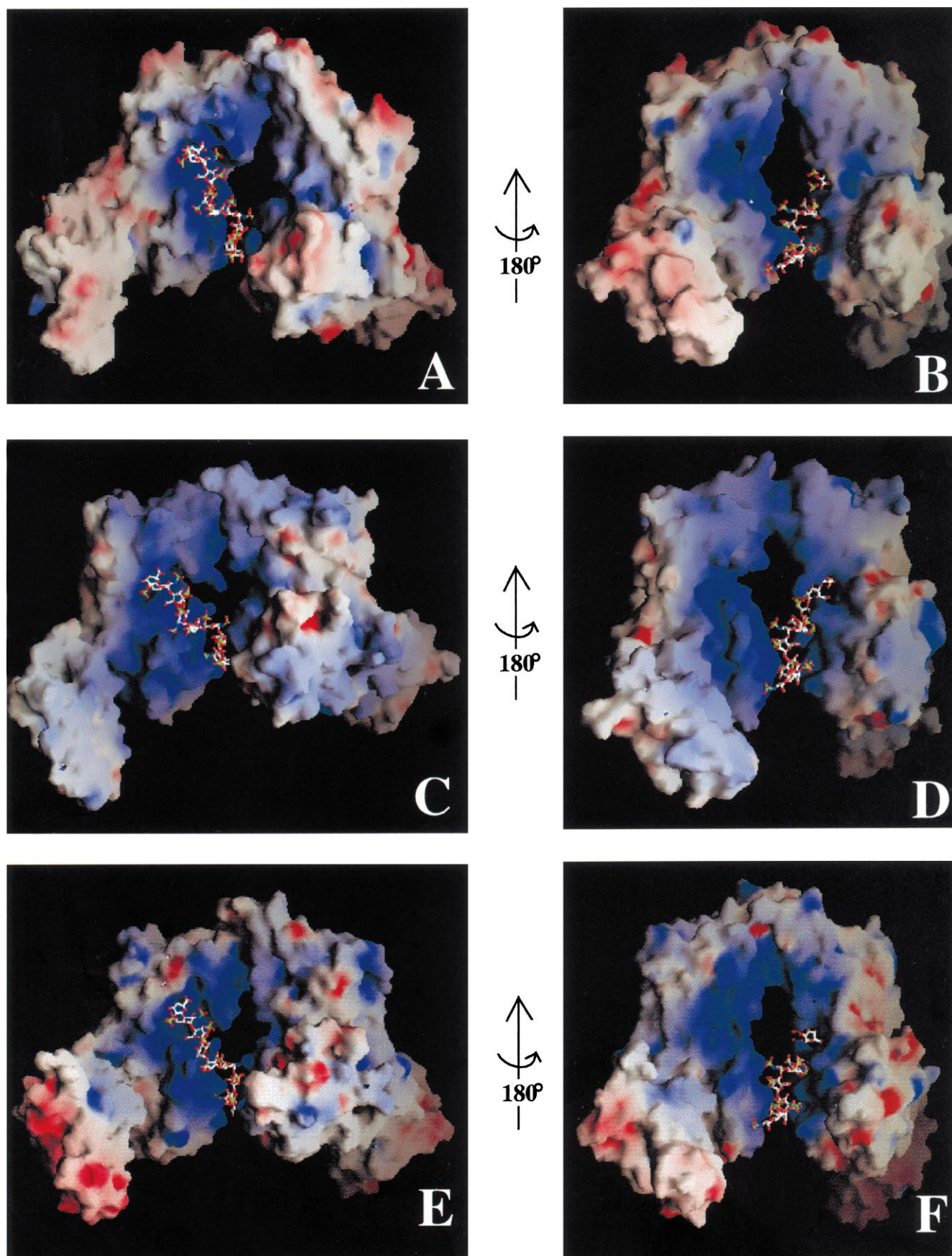
Fig. 3. Illustration of the charge profile of heparin-binding pocket in ternary complexes of 1e0o (A,B), Fly2 (C,D) and Worm1 (E,F). Heparin is depicted by rods. Images created by GRASP [71].

Worm2, respectively. Of these, 8 and 14 residues respectively are basic in nature. Fig. 2, generated by LIGPLOT [62], schematically illustrates these interactions between the basic residues and the heparin decasaccharide in the Fly1 and Worm2 models. Most dramatically, the overall charge distributions found in the human complexes are conserved in the fly and worm models (Fig. 3), especially in proximity to the heparin-binding pocket. This suggests that FGFs and FGFRs in *D. melanogaster* and *C. elegans* have binding sites for heparin similar to those observed in the human complexes.

Site-directed mutagenesis and other in vitro studies have identified specific residues that contribute significantly to heparin binding [21,31,63]. The interaction of a ligand residue, K125, in human FGF2 (its equivalent residue being K118 in 1e0o) is reported to be important in heparin binding as its mutation to an alanine leads to a 20% loss in binding energy [64]. Heparin interactions through the equivalent ligand residue in the fly (R109), let756 (R122) and egl17 (Q108) appear structurally maintained. Studies on the human receptor domains (FGFR1) have revealed an 18-residue peptide (K18K) that forms an obligatory heparin-binding domain [22], which corresponds to the amino acid stretch between K161 and R178 in 1e0o (Fig. 1). The heparin–receptor interactions in these equivalent stretches in *D. melanogaster* and *C. elegans* models also exhibit high similarity (Fig. 1 and Table 3).

Thus, the overall features of the interactions of FGF and FGFRs with heparin appear to be well conserved even in these primitive species, although the sequence identities with human homologues are as low as 30%. Table 3 shows that, although there are a number of substitutions of residues which interact with heparin in the human structure, there should be reasonable compensation in both the number and nature of interactions, thus preserving the overall affinity for heparin with species variation.

4. Conclusions

Comparative modeling studies on the *D. melanogaster* and *C. elegans* FGFs and their receptor sequences indicate that they can adopt similar conformations and make equivalent interactions to those observed in corresponding human structures. The analyses of heparin–protein contacts reveal a conserved binding surface for heparin in their FGFs and FGFRs. The similarity of the protein interface, particularly with respect to its charge distribution also suggests similar dimerisation and signaling mechanisms in these primitive organisms.

Whilst *D. melanogaster* and *C. elegans* maintain only a small number of FGFs and FGFRs, mammals exhibit a broad range of ligands and receptors. This may be related to the development of bone structures, and the highly developed immune and neuronal systems in vertebrates. Although there is no evidence for FGF involvement in immune development, multiple mutations have been documented which cause developmental defects in the skull, the long bones, and the digital bones [65–67]. Recently, several FGFs have been localised to specific subsets of neuronal tissues in the mouse brain [68–70]. These data support the hypothesis that FGFs are heavily involved in the correct formation of bone and complex neuronal structures, and it is possible that the increase in FGF numbers correlates with greater neuronal and skeletal complexity. It has been shown that some mutations found in Apert syndrome and other related disorders lead to an increase in receptor affinity and a broadening of specificity [57,58]. It is possible that fly and worm FGFRs developed similar features early in evolution, and that decreased affinity and increased specificity were a necessary prerequisite for subsequent diversification of roles of receptors and ligands.

References

[1] Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B. and Blundell, T.L. (2000) Nature 407, 1029–1034.
[2] Burgess, W.H. and Maciag, T. (1989) Annu. Rev. Biochem. 58, 575–606.
[3] Rifkin, D.B. and Moscatelli, D. (1989) J. Cell Biol. 109, 1–6.
[4] Baird, A. and Klagsbrun, M. (1991) Cancer Cells 3, 239–243.
[5] Yamaguchi, T.P. and Rossant, J. (1995) Curr. Opin. Genet. Dev. 5, 485–491.
[6] Martin, G.R. (1998) Genes Dev. 12, 1571–1586.
[7] Cuevas, P., Burgos, J. and Baird, A. (1988) Biochem. Biophys. Res. Commun. 156, 611–618.
[8] Clarke, M.S., Khakee, R. and McNeil, P.L. (1993) J. Cell Sci. 106, 121–133.
[9] Fernig, D.G. and Gallagher, J.T. (1994) Prog. Growth Factor Res. 5, 353–377.
[10] Powers, C.J., McLeskey, S.W. and Wellstein, A. (2000) Endocr.-Relat. Cancer 7, 165.
[11] Johnson, D.E. and Williams, L.T. (1993) Adv. Cancer Res. 60, 1–41.
[12] Bellot, F., Crumley, G., Kaplow, J.M., Schlessinger, J., Jaye, M. and Dionne, C.A. (1991) EMBO J. 10, 2849–2854.
[13] Wilkie, A.O.M., Morriss-Kay, G.M., Jones, E.Y. and Heath, J.K. (1995) Curr. Biol. 5, 500–507.
[14] Galzie, Z., Kinsella, A.R. and Smith, J.A. (1997) Biochem. Cell. Biol. 75, 669–685.
[15] Sutherland, D., Samakovlis, C. and Krasnow, M.A. (1996) Cell 87, 1091–1101.
[16] Lee, T., Hacohen, N., Krasnow, M., Montell, D.J., Steinberger, D., Reinhartz, T., Unsold, R. and Muller, U. (1996) Genes Dev. 10, 2912–2921.
[17] Beiman, M., Shilo, B.Z. and Volk, T. (1996) Genes Dev. 10, 2993–3002.
[18] Burdine, R.D., Branda, C.S. and Stern, M.J. (1998) Development 125, 1083–1093.
[19] Flaumenhaft, R., Moscatelli, D. and Rifkin, D.B. (1990) J. Cell. Biol. 111, 1651–1659.
[20] Lin, X.H., Buff, E.M., Perrimon, N. and Michelson, A.M. (1999) Development 126, 3715–3723.
[21] Rapraeger, A.C., Krufka, A. and Olwin, B.B. (1991) Science 252, 1705–1708.
[22] Kan, M., Wang, F., Xu, J., Crabb, J.W., Hou, J. and McKeehan, W.L. (1993) Science 259, 1918–1921.
[23] Mach, H., Volkin, D.B., Burke, C.J., Middaugh, C.R., Linhardt, R.J., Fromm, J.R. and Loganathan, D. (1993) Biochemistry 32, 5480–5489.
[24] Spivak-Kroizman, T. (1994) Cell 79, 1015–1024.
[25] Gao, G. and Goldfarb, M. (1995) EMBO J. 14, 2183–2190.
[26] Faham, S., Linhardt, R.J. and Rees, D.C. (1998) Curr. Opin. Struct. Biol. 8, 578–586.
[27] Satoh, H., Susaki, M., Shukunami, C., Iyama, K., Negoro, T. and Hiraki, Y. (1998) J. Biol. Chem. 273, 12307–12315.
[28] Waksman, G. and Herr, A.B. (1998) Nat. Struct. Biol. 5, 527–530.
[29] Pye, D.A. and Gallagher, J.T. (1999) J. Biol. Chem. 274, 13456–13461.
[30] Hsu, Y.R., Nybo, R., Sullivan, J.K., Costigan, V., Spahr, C.S., Wong, C., Jones, M., Pentzer, A.G., Crouse, J.A., Pacifici, R.E., Lu, H.S., Morris, C.F. and Philo, J.S. (1999) Biochemistry 38, 2523–2534.
[31] Baeg, G. and Perrimon, N. (2000) Curr. Opin. Cell Biol. 12, 575–580.
[32] Ornitz, D.M. (2000) Bioessays 22, 108–112.

[33] Roubin, R., Naert, K., Popovici, C., Vatcher, G., Coulier, F., Thierry-Mieg, J., Pontarotti, P., Birnbaum, D., Baillie, D. and Thierry-Mieg, D. (1999) Oncogene 18, 6741–6747.

[34] DeVore, D.L., Horvitz, H.R. and Stern, M.J. (1995) Cell 83, 611–620.

[35] Burdine, R.D., Chen, E.B., Kwok, S.F. and Stern, M.J. (1997) Proc. Natl. Acad. Sci. USA 94, 2433–2437.

[36] Shishido, E., Higashijima, S., Emori, Y. and Saigo, K. (1993) Development 117, 751–761.

[37] Schlessinger, J., Plotnikov, A.N., Ibrahami, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A., Linhardt, R.J. and Mohammadi, M. (2000) Mol. Cell 6, 743–750.

[38] Shi, J., Blundell, T.L. and Mizuguchi, K. (2001), communicated.

[39] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.

[40] Galtier, N., Guoy, M. and Gautier, C. (1996) Comput. Appl. Biosci. 12, 543–548.

[41] Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. (1998) Bioinformatics 14, 617–623.

[42] Sali, A. and Sánchez, R. (1999) J. Comp. Phys. 151, 388–401.

[43] Sali, A. and Blundell, T.L. (1993) J. Mol. Biol. 234, 779–815.

[44] Deane, C.M. and Blundell, T.L. (2000) Proteins 40, 135–144.

[45] Bower, M.J., Cohen, F.E. and Dunbrack, R.L. (1997) J. Mol. Biol. 267, 1268–1282.

[46] Jones, T.A., Xhou, J.Y. and Cowan, S.W. (1991) Acta Crystallogr. A47, 110–119.

[47] Guex, N. and Peitsch, M.C. (1997) Electrophoresis 18, 2714–2723.

[48] Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. and Karplus, M. (1995) Proteins 23, 318–326.

[49] Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993) J. Mol. Biol. 231, 1049–1067.

[50] Luthy, H., Bowie, J.U. and Eisenberg, D. (1992) Nature 356, 83–85.

[51] Collaborative Computational Project Number 4, (1994) Acta Crystallogr. D50, 760–763.

[52] Coulier, F., Pontarotti, P., Roubin, R., Hartung, H., Goldfarb, M. and Birnbaum, D. (1997) J. Mol. Evol. 44, 43–56.

[53] Stauber, D.J., DiGabriele, A.D. and Hendrikson, W.A. (2000) Proc. Natl. Acad. Sci. USA 97, 49–54.

[54] Zhu, H., Anchin, J., Ramnarayan, K., Zheng, J., Kawai, T., Mong, S. and Wolff, M.E. (1997) Protein Eng. 10, 417–421.

[55] Zhu, H., Ramnarayan, K., Anchin, J., Miao, W.Y., Sereno, A., Millman, L., Zheng, J., Balaji, V.N. and Wolff, M.E. (1995) J. Biol. Chem. 270, 21869–21874.

[56] Cohen Jr., M.M. and Kreiborg, S. (1990) Am. J. Med. Genet. 135, 36–45.

[57] Anderson, J., Burns, H.D., Enriquez-Harris, P., Wilkie, A.O.M. and Heath, J.K. (1998) Hum. Mol. Genet. 7, 1475–1483.

[58] Yu, K., Herr, A.B., Waksman, G. and Ornitz, D.M. (2000) Proc. Natl. Acad. Sci. USA 97, 14536–14541.

[59] Muenke, M., Schell, U., Hehr, H., Robin, N.H., Losken, H.W., Schinzel, A., Pullyen, L.J., Rutland, P., Reardon, W. and Malcolm, S. (1994) Nat. Genet. 8, 269–274.

[60] Bellus, G.A., Gaudenz, K., Zackai, E.H., Clarke, L.A., Szabo, J., Francomano, C.A. and Muenke, M. (1996) Nat. Genet. 14, 174–176.

[61] Wang, F., Kan, M., Xu, J., Yan, G. and McKeehan, W.L. (1995) J. Biol. Chem. 270, 10222–10230.

[62] Wallace, A.C., Laskowski, R.A. and Thornton, J.M. (1995) Protein Eng. 8, 127–134.

[63] Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G.M., Arakawa, T., Hsu, B.T. and Rees, D.C. (1991) Science 251, 90–93.

[64] Thompson, L.D., Pantoliano, M.W. and Springer, B.A. (1994) Biochemistry 33, 3831–3840.

[65] Wilkie, A.O.M. (1997) Hum. Mol. Genet. 6, 1647–1656.

[66] Burke, D., Wilkes, D., Blundell, T.L. and Malcolm, S. (1998) Trends Biochem. Sci. 23, 59–63.

[67] Webster, M.K. and Donoghue, D.J. (1997) Trends Genet. 13, 178–182.

[68] Yamashita, T., Yoshioka, M. and Itoh, N. (2000) Biochem. Biophys. Res. Commun. 277, 494–498.

[69] Ohmachi, S., Watanabe, Y., Mikami, T., Kusu, N., Ibi, T., Akaike, A. and Itoh, N. (2000) Biochem. Biophys. Res. Commun. 277, 355–360.

[70] Ohuchi, H., Kimura, S., Watamoto, M. and Itoh, N. (2000) Mech. Dev. 95, 55–66.

[71] Nicholls, A., Sharp, K. and Honig, B. (1991) Proteins 11, 281–296.