Insights into the molecular basis for fibroblast growth factor receptor autoinhibition and ligand-binding promiscuity

Shaun K. Olsen*, Omar A. Ibrahimi*, Angela Raucci†, Fuming Zhang‡, Anna V. Eliseenkova*, Avner Yayon§, Claudio Basilico†, Robert J. Linhardt‡, Joseph Schlessinger¶, and Moosa Mohammadi*‖

Departments of *Pharmacology and †Microbiology, New York University School of Medicine, New York, NY 10016; ‡Department of Chemistry, Biology, and Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180; ¶Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520; and §ProChon Biotech Ltd., Weizmann Science Park, P.O. Box 1482, Rehovot 76114, Israel

Contributed by Joseph Schlessinger, November 26, 2003

The prototypical fibroblast growth factor receptor (FGFR) extracellular domain consists of three Ig domains (D1–D3) of which the two membrane-proximal D2 and D3 domains and the interconnecting D2–D3 linker bear the determinants of ligand binding and specificity. In contrast, D1 and the D1–D2 linker are thought to play autoinhibitory roles in FGFR regulation. Here, we report the crystal structure of the three-Ig form of FGFR3c in complex with FGF1, an FGF that binds promiscuously to each of the seven principal FGFRs. In this structure, D1 and the D1–D2 linker are completely disordered, demonstrating that these regions are dispensable for FGF binding. Real-time binding experiments using surface plasmon resonance show that relative to two-Ig form, the three-Ig form of FGFR3c exhibits lower affinity for both FGF1 and heparin. Importantly, we demonstrate that this autoinhibition is mediated by intramolecular interactions of D1 and the D1–D2 linker with the minimal FGF and heparin-binding D2–D3 region. As in the FGF1–FGFR2c structure, but not the FGF1–FGFR1c structure, the alternatively spliced βC′–βE loop is ordered and interacts with FGF1 in the FGF1–FGFR3c structure. However, in contrast to the FGF1–FGFR2c structure in which the βC′–βE loop interacts with the β-trefoil core region of FGF1, in the FGF1–FGFR3c structure, this loop interacts extensively with the N-terminal region of FGF1, underscoring the importance of the FGF1 N terminus in conferring receptor-binding affinity and promiscuity. Importantly, comparison of the three FGF1–FGFR structures shows that the flexibility of the βC′–βE loop is a major determinant of ligand-binding specificity and promiscuity.

The mammalian fibroblast growth factors (FGFs) (FGF1–FGF23) constitute a large family of structurally related ligands that are involved in a multitude of biological processes from embryogenesis to adult homeostasis (1). The biological effects of FGFs are mediated by binding to the FGF receptor (FGFR) family of receptor tyrosine kinases (FGFR1–FGFR4). The prototypical FGFR consists of an extracellular domain composed of three Ig domains (D1, D2, and D3), a single-pass transmembrane helix, and a cytoplasmic tyrosine kinase domain. A contiguous stretch of 4–8 acidic amino acids within the D1–D2 linker has been termed the acid box (2–4). FGFs have been shown to require heparan sulfate or soluble heparin to bind with high affinity to FGFRs and to induce receptor dimerization and activation (5–7).

The regulation of ligand-binding specificity of FGFR is essential for the control of FGF signaling and is primarily achieved by alternative splicing of FGFR. An alternative splicing event in FGFR1–3 involving the exon encoding the C-terminal region of D3 results in two otherwise identical receptors (b and c isoforms), possessing different ligand-binding specificities (8–10). Whereas most FGFs activate a particular subset of FGFRs, FGF1 activates each of the seven principal FGFRs, and, as such, has been termed the ''universal'' ligand (11). Additional alternative splicing events result in FGFR isoforms that lack D1 (in FGFR1 and 2), D1 and the D1–D2 linker (in FGFR2), or the D1–D2 linker (in FGFR3; refs. 3, 4, and 12). The biological significance of alternative splicing of D1 and the D1–D2 linker is not fully characterized. The recent crystal structures of several FGFs in complex with fragments of FGFRs encompassing D2 and D3 demonstrate that D2, D3, and D2–D3 linker of FGFRs are necessary and sufficient for specific ligand binding (13–16). In addition, comparison of the FGF2–FGFR2c, and FGF10–FGFR2b structures have revealed that specific contacts made with the alternatively spliced βC′–βE loop in D3 are the principal determinants of ligand-binding specificity (14, 16).

In contrast, a study in which chimeric FGFR1 and FGFR3 constructs were used to map the determinants of ligand binding and specificity suggested that the D1–D2 linker of FGFR3c confers specific binding to FGF8 and FGF9 (17). Reciprocally, studies by Mckeehan and coworkers (18) suggest that, although D1 and the D1–D2 linker are dispensable for ligand binding and affinity, these regions play an autoregulatory role in FGFR function. These studies show that the three-Ig form of FGFR1c has a lower affinity for heparin and FGF1 than the two-Ig form, suggesting that D1 autoinhibits ligand- and heparin-binding affinity of FGFR (18). Because deletion of the D1–D2 linker was sufficient to relieve this inhibition, it was proposed that the D1–D2 linker region serves as a flexible hinge that allows D1 to interact with the D2–D3 portion of FGFR to alter affinity for ligand and heparin (18).

To investigate the role of D1 and the D1–D2 linker in FGFR regulation, we solved the crystal structure of the complete extracellular domain of FGFR3c in complex with FGF1. In the FGF1–FGFR3c structure, D1 and the D1–D2 linker are disordered, demonstrating that only D2, D3, and the D2–D3 linker participate in FGF1 binding. In addition, comparison of this structure to the other FGF1–FGFR structures shows that the promiscuity of FGF1 is a result of its ability to bind FGFRs with high affinity in the presence or absence of contacts with the alternatively spliced βC′–βE loop in D3. Importantly, the structural and biochemical data substantiates a model in which D1 and the acid box act cooperatively to negatively regulate FGFR function by competing with FGF and heparin for FGFR binding.

Methods

Protein Purification. Full-length FGF1 (residues 1–155), and the N-terminally truncated FGF1 (residues 21–155) were expressed in *Escherichia coli* and purified by heparin affinity and cation exchange chromatography. The three-Ig form of human FGFR3c (residues 33–365), and the two-Ig form (residues

Abbreviations: FGF, fibroblast growth factor; FGFR, FGF receptor; RCS, rat chondrosarcoma; SPR, surface plasmon resonance.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1RY7).

‖To whom correspondence should be addressed. E-mail: mohammad@saturn.med.nyu.edu.

143–365) were expressed in *E. coli* and were refolded from inclusion bodies by using a described protocol (14). Properly folded three- and two-Ig forms of FGFR3c were purified by heparin affinity, size-exclusion, and anion-exchange chromatography. The FGFR3c fragment consisting of D1 plus the D1–D2 linker (D1+; residues 33–145) was expressed in *E. coli*. Soluble D1+ was purified by using anion-exchange, nickel affinity, and size-exclusion chromatography.

Crystallization, Structure Determination, and Refinement. Crystals were grown at 20 C° by using hanging drop vapor diffusion. Crystals of the FGF1–FGFR3c complex were obtained by mixing 1.5 μl of crystallization buffer (16% polyethylene glycol 4000/0.2 M ammonium sulfate/5% glycerol/5% 2-methyl-2,4-pentandiol/25 μM cadmium chloride) with 1.5 μl of protein solution (14 mg/ml/25 mM Hepes, pH 7.5/150 mM NaCl). The FGF1–FGFR3c crystals are in the C-centered monoclinic space group C2, with $a = 104.22$ Å, $b = 64.75$ Å, $c = 99.93$ Å, and $\beta = 94.67°$. The asymmetric unit contains a single FGF1–FGFR3c complex. Crystals were flash-frozen in a final cryoprotectant solution composed of 8% glycerol, 8% ethylene glycol, and 8% 2-methyl-2,4-pentandiol. Diffraction data for the FGF1–FGFR3c complex was measured on a Q4R charge-coupled device detector at beamline X4A at the National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY. The data were processed with DENZO and SCALEPACK (19).

The program AMORE (20) was used to find a molecular replacement solution by using the structure of the FGF1–FGFR1c complex (PDB ID code 1EVT; ref. 14) as the search model. Rigid body, positional, and B-factor refinement, and simulated annealing were performed by using CNS (21). The program o (22) was used for model building into the $2F_o - F_c$ and $F_o - F_c$ maps. The refined model for the FGF1–FGFR3c complex is composed of one FGF1 molecule (residues 5–155) and one FGFR3c molecule (residues 150–362). In this structure, D1 and the D1–D2 linker of FGFR3c (residues 33–149) are disordered. Residues 1–4 of FGF1 are also disordered.

Surface Plasmon Resonance (SPR) Analysis. Real-time biomolecular interactions were analyzed with a BIAcore 3000 system (Biosensor AB, Uppsala). The two-Ig form of FGFR3c, and full-length and truncated FGF1, were immobilized on research grade CM5 chips (Biosensor AB) according to a previously reported protocol (23). A heparin neoproteoglycan chip was prepared as reported (23). Full-length FGF1 was immobilized to ≈1,000 response units (RUs). Truncated FGF1 and the two-Ig form of FGFR3c were immobilized to corresponding RU levels based on their molecular masses. To obtain kinetic data, different concentrations of analytes in HBS-EP buffer (0.01 M Hepes/0.15 M NaCl/3 mM EDTA/0.005% polysorbate 20, pH 7.4) were injected over the protein sensor chip at a flow rate of 50 μl/min. At the end of each sample injection (180 s), HBS-EP buffer was passed over the sensor surface to monitor the dissociation phase. Kinetic parameters for each interaction, with the exception of the interaction between D1+ and the two-Ig form of FGFR3c, was determined by globally fitting the experimental data to a 1:1 interaction with BIAEVALUATION software (Biosensor AB). The dissociation constant of the interaction between D1+ and the two-Ig form of FGFR3c was determined by globally fitting the experimental data to a 1:1 steady-state interaction. A minimum of four different analyte concentrations were used to determine the kinetic parameters for each interaction.

Immunoprecipitation and Immunoblot Analysis. Rat chondrosarcoma (RCS) cells were maintained as described (24). RCS cells were starved overnight in DMEM and were stimulated for 5 min at 37°C with the indicated concentration of truncated or full-length FGF1 in DMEM. Because RCS cells endogenously express heparan sulfate proteoglycan, soluble heparin was not added. Cell lysis, immunoprecipitation, and Western blotting were carried out as described (24).

Table 1. Summary of crystallographic analysis

Data collection statistics	
Resolution, Å	30.0–3.2
Reflections (total/unique)	41,929/11,404
Completeness, %	98.9 (95.5)†
R_{sym} %*	6.8 (30.3)†
Signal ($\langle I/\sigma I \rangle$)	13.8
Refinement statistics‡	
Resolution, Å	25.0–3.2
Reflections	10,502
R_{cryst}/R_{free}, %§	27.7/34.1
rms deviations	
Bonds, Å	0.008
Angles, °	1.5
B factors, Å²¶	2.3

*$R_{sym} = 100 \times \Sigma_{hkl}\Sigma|I_i(hkl) - |\langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$.
†Value in parentheses is for the highest-resolution shell: 3.31–3.2 Å.
‡Atomic model: 2,768 protein atoms.
§$R_{cryst/free} = 100 \times \Sigma_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \Sigma_{hkl} |F_o(hkl)|$, where F_o ($>0\sigma$) and F_c are the observed and calculated structure factors, respectively. Ten percent of the reflections were used for calculation of R_{free}.
¶For bonded protein atoms.

Results and Discussion

For previous crystallographic studies of FGF1–FGFR complexes, we and others used an N-terminally truncated FGF1 (residues 21–155) that binds with high affinity to FGFR1c and FGFR2c (14, 15). Interestingly, the truncated FGF1 construct bound poorly to full-length FGFR3c on a size-exclusion column, whereas full-length FGF1 formed a stable 1:1 complex (data not shown). The crystal structure of full-length FGF1 in complex with the three-Ig domain form of FGFR3c was solved by molecular replacement by using the two-Ig domain containing the FGFR1c–FGF1 structure as the search model. After multiple rounds of refinement, there was clear electron density for FGF1 and the D2–D3 region of FGFR3c, but the density for D1 and the D1–D2 linker was absent. We believe that the flexibility of the D1–D2 linker results in a random orientation of D1 relative to the rest of the structure, thereby accounting for the lack of interpretable electron density. Indeed, analysis of crystal packing shows that there is sufficient unoccupied space in the vicinity of D2 within the unit cell that could accommodate D1. The FGF1–FGFR3c structure has been refined to 3.2 Å with working and free *R* values of 27.7% and 34.1%, respectively (Table 1). The modest working and free *R* factors for the FGF1–FGFR3c structure are likely the result of our inability to trace D1 and the D1–D2 linker. A ribbon representation of the equivalent of a two-Ig domain form of FGFR3c in complex with FGF1 is presented in Fig. 1*A*.

To confirm that D1 and the D1–D2 linker were present in the crystal, several crystals representative of those used for data collection were dissolved and analyzed by using SDS/PAGE. Protein from dissolved FGF1–FGFR3c crystals was virtually indistinguishable from freshly purified complex, making it unlikely that a proteolytic event facilitated crystallization during incubation at 20 C° (Fig. 1*A Inset*). Therefore, the FGF1–FGFR3c crystal structure demonstrates that D1 and the acid box do not participate in FGF1 binding.

Structural Basis for the Promiscuity of FGF1. Because of its ability to activate each of the seven principal FGFRs in living cells, FGF1 has been considered to be the universal FGFR ligand (11). Although the crystal structures of FGF1 in complex with

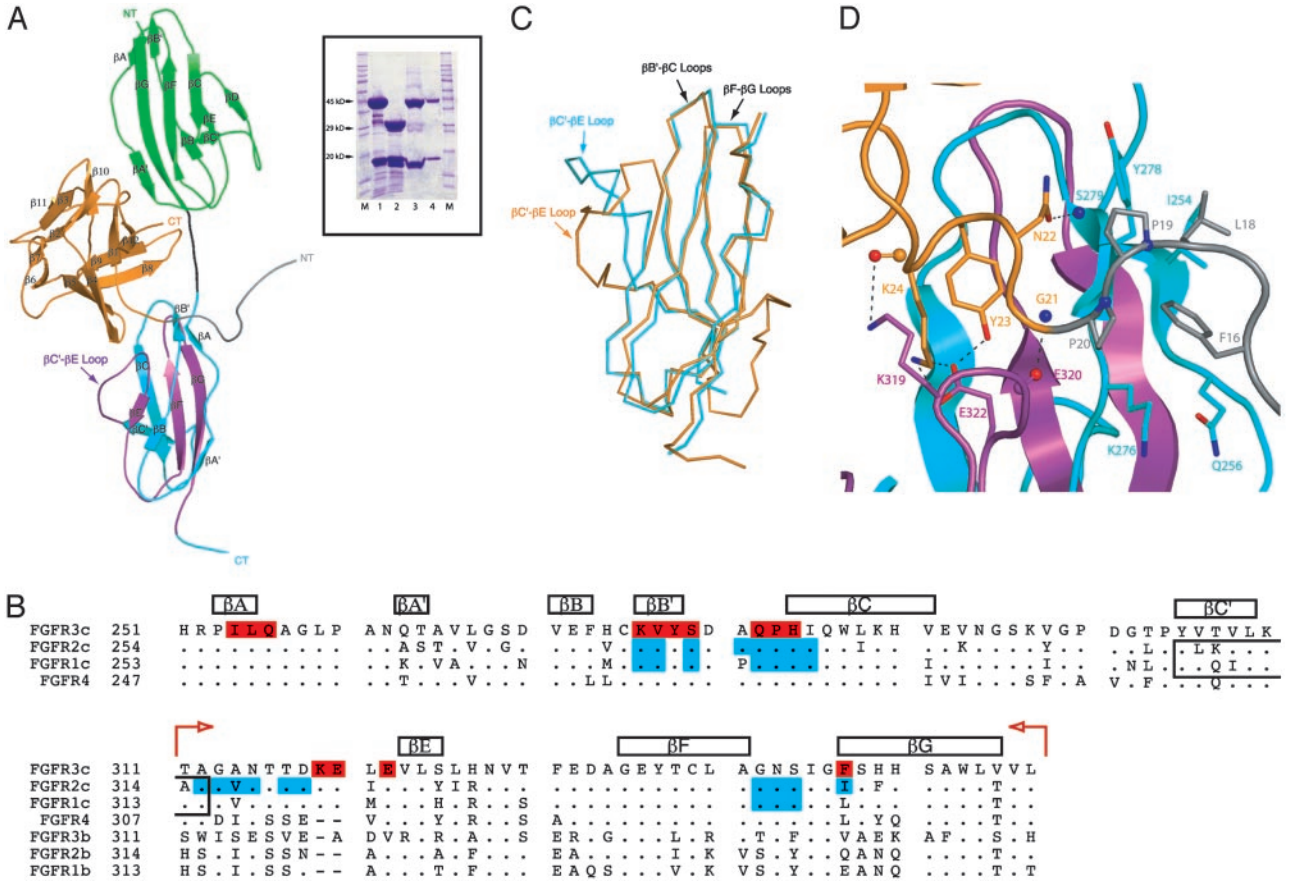

Fig. 1. The FGFR3c–FGF1 structure. (*A*) A ribbon representation of the FGFR3c–FGF1 complex. FGF1 is orange, FGFR3c D2 is green, D3 is cyan, and the D2–D3 linker is black. The alternatively spliced C-terminal half of D3 is purple. The N and C termini of FGF1 are labeled NT and CT, respectively. The FGF1 N-terminal region not included in the truncated FGF1 construct used in the FGFR1c and FGFR2c structures is gray. (*Inset*) Protein from FGF1–FGFR3c crystals is indistinguishable from the freshly purified complex. SDS/PAGE analysis of the purified three-Ig form FGFR3c–FGF1 complex (1), the purified two-Ig form FGFR3c–FGF1 complex (2), the protein solution from a FGF1–FGFR3c crystal containing hanging drop (3), and dissolved FGF1–FGFR3c crystals (4) (washed twice). Lanes with molecular mass markers are labeled M, and selected molecular masses are labeled. (*B*) Structure-based sequence alignment of Ig domain 3 from human FGFRs. The sequence alignment was performed by using CLUSTALW (28). The location and length of the β strands are shown on top of the sequence alignment. Note that the βC′ strand of FGFR3c terminates two residues earlier than those of FGFR1c and FGFR2c, and, therefore, the βC′–βE loop of FGFR3 is 12 residues long (residues 310–322). The different lengths of the βC′–βE loops of FGFR1c and FGFR2c from the FGF1–FGFR structures are indicated by boxes within the alignment. A period indicates sequence identity to FGFR3c. A dash represents a gap introduced to optimize the alignment. The alternatively spliced C-terminal half of D3 is marked by red arrows. FGFR3c residues that interact with FGF1 are red. FGFR1c and FGFR2c residues that interact with FGF in other crystal structures are cyan. (*C*) Superimposition of FGFR3c and FGFR2c D3. The C^{α} trace of FGFR2c D3 from the FGFR2c–FGF1 structure (cyan) was superimposed onto the C^{α} trace of FGFR3c D3 from the FGFR3c–FGF1 structure (orange; rms deviation = 0.831 Å). Residues corresponding to the βC–βC′ and βC′–βE loops were not included in the superimposition. The βB′–βC, βC′–βE, and βF–βG loops of the D3s are marked by an arrowhead. (*D*) Interactions between the FGF1 N terminus and D3 in the FGFR3c–FGF1 structure. Colors are as in *A*. Selected residues are labeled and are rendered in a stick format. This figure was created by using the program PYMOL (29).

FGFR1c and FGFR2c have been reported, the structural basis for the promiscuity of FGF1 is not fully understood. In the FGF1–FGFR1c structure, FGF1 does not interact with the alternatively spliced βC′–βE loop (residues 315–324), and, as a result, this loop is completely disordered. The lack of specific contacts with the βC′–βE loop was proposed to account for why FGF1 binds to all FGFRs indiscriminate of alternative splicing (14). In contrast, the βC′–βE loop (residues 316–325) in the FGF1–FGFR2c structure is ordered and interacts with FGF1. Most notably, Val-317 from the βC′–βE loop engages in specific hydrophobic contacts and hydrogen bond with residues within the β-trefoil core of FGF1 (15). Therefore, it is unclear how the promiscuity of FGF1 is conferred at the molecular level. Analysis of the FGF1–FGFR3c structure provides insights into the molecular mechanism underlying the promiscuity of FGF1.

The overall FGF1–FGFR3c structure is similar to the previously determined FGF1–FGFR structures. The only major differences between the FGF1–FGFR3c structure and the other FGF1–FGFR structures are the βC′–βE loop conformation and the accompanying interface between this loop and FGF1. As in the FGF1–FGFR2c structure, the βC′–βE loop in FGF1–FGFR3c is ordered. However, the conformation of the loop is drastically different from that of FGFR2c (Fig. 1*C*). This finding is surprising, given the fact that the βC′–βE loop of FGFR3c differs only by two and three conservative amino acid changes from the corresponding loops of FGFR1c and FGFR2c, respectively (Fig. 1*B*). Comparison of the FGF1–FGFR3c structure to the other two structures suggests these changes facilitate the altered path of the βC′–βE loop in FGFR3c, which results in a completely different set of interactions between FGF1 and the loop (Fig. 1*B*).

In both the FGF1–FGFR3c and FGF1–FGFR2c structures, the conformation of the βC′–βE loop is primarily dictated by the contacts the loop makes with the ligand. In the FGF1–FGFR2c structure, the hydrophobic contact between Val-317 from the loop and residues within the β-trefoil core region of FGF1 facilitates the path of the βC′–βE loop. In the FGF1–FGFR3c structure, this

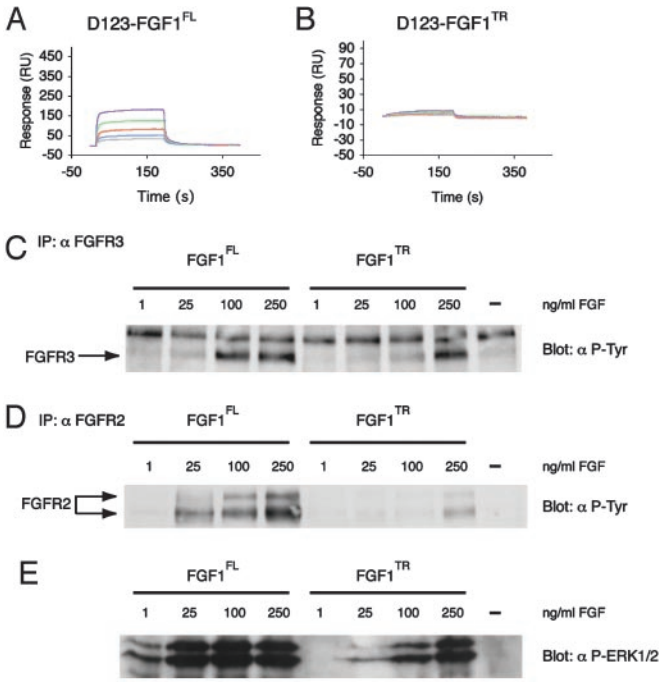

Fig. 2. The affinity and activation of FGFRs by FGF1 is enhanced by the N terminus. (*A* and *B*) Sensorgrams of the full-length versus truncated FGF1 interaction with the D123 isoform of FGFR3c. Analyte concentrations are colored as follows: purple (1.6 μM), green (0.8 μM), red (0.4 μM), blue (0.2 μM), and gray (0.1 μM). (*C–E*) Full-length FGF1 activates FGFR3 and FGFR2 more potently than truncated FGF1 on RCS cells (*C–D*). Analysis of FGFR3 and FGFR2 phosphorylation. Clarified cell lysates of RCSs starved overnight and treated with the indicated concentration of FGF for 5 min were examined by using protein immunoblot. (*E*) FGFR3 and FGFR2 immunoblots were probed with a phosphotyrosine-specific antibody. Analysis of ERK1/2 phosphorylation. Clarified cell lysates used in *A* were probed with an antibody specific for anti-phospho-ERK1/2.

Table 2. Summary of kinetic data

	FGFR3c	
FGF/heparin	D2–D3	D1–D3
FL-FGF1		
k_{on}*	8.79×10^{5}	1.96×10^{5}
k_{off}*	2.02×10^{-1}	1.8×10^{-1}
K_d†	2.30×10^{-7}	9.16×10^{-2}
TR-FGF1		
k_{on}	28.0	NB
k_{off}	1.9×10^{-2}	NB
K_d	6.79×10^{-4}	NB
Heparin		
k_{on}	4.75×10^{4}	3.44×10^{3}
k_{off}	0.0374	0.0155
K_d	7.88×10^{-7}	4.50×10^{-6}
D1+		
k_{on}	—	—
k_{off}	—	—
K_d	2.0×10^{-5}	—

NB, negligible binding.

*k_{on} (per mol/s) and k_{off} (per s) were derived as described in *Methods*. χ^2 was $<10\%$ in all cases.

†The apparent affinity, K_d (per mol), is equal to k_{off}/k_{on}.

hydrophobic interaction with the FGF1 core region does not take place. This is because Val-317 is an alanine (Ala-314) in FGFR3c and has weaker propensity to interact hydrophobically with FGF1. Because of the loss of this hydrophobic contact, the βC′–βE loop of FGFR3c moves away from the β-trefoil core region of FGF1 and instead interacts with the N-terminal region of FGF1 outside of the β-trefoil core. Specifically, FGFR-invariant Lys-319, Glu-320, and Glu-322 at the C terminus of the βC′–βE loop engage in number of unique hydrogen bonds with Tyr-23 and Lys-24 from the FGF1 N terminus (Fig. 1*D*). Moreover, a backbone-mediated hydrogen bond occurs between Glu-320 of the loop with Gly-21 of the N terminus (Fig. 1*D*). These specific interactions enhance FGF1–FGFR3c affinity and stabilize the observed path of βC′–βE loop in FGFR3c.

Whereas contacts with ligand are the major facilitators of the βC′–βE loop conformation, intraloop interactions and interactions between the loop and the main body of D3 also contribute to loop stability. However, these intramolecular interactions are drastically different between the two structures. In FGFR2c, Ala-315 and Ile-324 from the βC′–βE loop form a hydrophobic patch with the FGFR-invariant Ile-288 in the βC strand, which tethers the loop to the main body of D3 and stabilizes the loop conformation. Ala-315 is conserved in FGFR3c (Ala-312), and Ile-324 is conservatively substituted by Leu-321 in FGFR3c (Fig. 1*B*). Surprisingly, these FGFR3c residues do not form the predicted hydrophobic patch with Ile-285 (corresponding with Ile-288 of FGFR2c). Instead, these residues tether the βC′–βE loop to the main body of D3 through a different mechanism. Leu-321 engages in a hydrophobic contact with the aliphatic portion of the Lys-276 side chain in the βB′ strand. The backbone of Ala-312 makes a hydrogen bond with Ser-325 from the βE strand (Fig. 1*B*), a residue that is unique to FGFR3c (Fig. 1*B*). Taken together, both ligand receptor as well as intrareceptor interactions combine to determine the observed conformation of the βC′–βE loop in each FGF–FGFR structure.

Flexibility of the βC′–βE Loop Modulates Ligand-Binding Specificity/Promiscuity. We surmise that, in the absence of FGF ligand, the intrareceptor interactions are not sufficient to confer a stable fold to the βC′–βE loop and that, therefore, the βC′–βE loop is probably flexible. Importantly, because the nature and extent of the intramolecular interactions facilitating the conformation of the βC′–βE loop are different in each structure, we predict that subtle variations exist in the loop flexibility of unliganded receptors. The inherent flexibility of the loop implies that interactions between FGF and the βC′–βE loop are associated with entropy loss. For a net gain in binding affinity, the sum of the interactions that FGF makes with the loop must outweigh the entropy cost associated with binding. These differences in loop flexibility differentially promote interactions of FGF with the βC′–βE loop and thereby modulate FGF–FGFR affinity.

This proposal provides a satisfying explanation for why the βC′–βE loop is disordered in the FGF1–FGFR1c crystal structure despite the fact that Val-317 of FGFR2c is conserved in FGFR1c (Val-316; Fig. 1*B*). Namely, the residue corresponding to Ile-324 is Met-323, and has a lower propensity to participate in the formation of the hydrophobic patch with Ala-314 and Ile-287 (residues corresponding to Ala-315 and Ile-288 of FGFR2c). This substitution weakens the tethering of the βC′–βE loop to the main body of D3, thereby increasing the flexibility of the loop in FGFR1c. This increased flexibility discourages the analogous Val-316–FGF interaction and thereby results in a disordered βC′–βE loop in the FGF1–FGFR1c structure. The plasticity of the βC′–βE loop is harmonious with this loop playing a decisive role in both ligand-binding specificity and promiscuity.

In summary, the promiscuity of FGF1 is the result of its unique ability to differentially interact with the inherently flexible βC′–βE loop. Specifically, FGF1 binding to FGFR can occur in the presence or absence of contacts with the alternatively spliced βC′–βE loop. This result accounts for why, despite the fact that

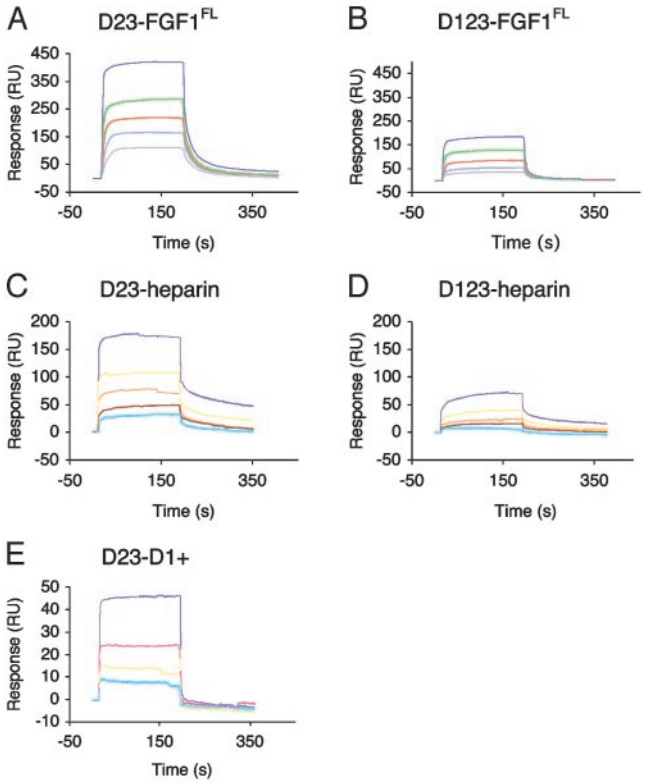

Fig. 3. D1 and the D1–D2 linker negatively regulate ligand- and heparin-binding affinity of FGFR3c. (*A* and *B*) Sensorgrams of the two- versus three-Ig form of FGFR3c binding to full-length FGF1. Analyte concentrations are colored as follows: purple (1.6 μM), green (0.8 μM), red (0.4 μM), blue (0.2 μM), and gray (0.1 μM). (*C* and *D*) Sensorgrams of the two- versus three-Ig form of FGFR3c binding to heparin. Analyte concentrations are colored as follows: purple (5 μM), yellow (2.5 μM), orange (1.25 μM), red (0.625 μM), and cyan (0.313 μM). (*E*) Sensorgram of D1+ binding to the two-Ig form of FGFR3c. Analyte concentrations are colored as follows: purple (10 μM), pink (5 μM), yellow (2.5 μM), and cyan (1 μM).

the primary sequence of the βC′–βE loop is highly similar between three receptors, the loop adopts drastically different conformations in each structure. Given the differences in loop flexibility and the magnitude of interactions that FGF1 makes with this loop in each of the three complexes, we predict that there are subtle differences in the affinity of FGF1 for FGFR1c, FGFR2c, and FGFR3c. Furthermore, because most FGFs activate a subset of FGFRs, it is likely that variations of the structural mechanism by which FGF1 attains promiscuity applies to other FGFs.

The FGF1 N terminus Is Required for High-Affinity Binding and Activation of FGFR3c. FGF N termini, defined structurally as the region preceding the globular β-trefoil core homology domain, are inherently flexible, as evident from their disordering in the crystal structures of unbound FGFs (25). Because protein flexibility can be a major obstacle in protein crystallization, we and others previously used an N-terminally truncated FGF1 (residues 21–155) to generate complexes with FGFR1c and FGFR2c (14, 15). This truncated ligand binds well to both FGFR1c and FGFR2c as determined by our ability to isolate a stable 1:1 complex on the size-exclusion column. In both of these structures, the few remaining N-terminal residues (21–26) are ordered as several residues from this remaining region interact with FGFR D3. Interestingly, we were able to isolate a stable complex of only full-length FGF1 with FGFR3c, because truncated FGF1 bound poorly to FGFR3c.

Nearly the entire FGF1 N terminus is ordered in the FGF1–FGFR3c crystal structure. The ordering of residues 5–15 is due to favorable crystal lattice contacts, whereas the ordering of the remaining residues (16–26) results from the interactions of this region with receptor D3. Importantly, Phe-16, Leu-18, Pro-19, and Pro-20 of FGF1, not present in the truncated FGF1, engage in a hydrophobic contact with Ile-254, Gln-256, Lys-276, and Tyr-278 on the surface of D3 (Fig. 1*D*). These additional interactions provide a plausible explanation for why we were only able to isolate a complex of full-length FGF1 with FGFR3c by using size exclusion chromatography. As in the previous FGF1–FGFR structures, the remaining N-terminal residues (21–26) of FGF1 also interact receptor D3 in the FGF1–FGFR3c structure. However, in the FGF1–FGFR3c structure, these residues primarily interact with the βC′–βE loop (Fig. 1*D*), whereas in the other two structures, these residues interact with the first half of D3. Moreover, interactions between the N terminus and D3 are more extensive in the FGF1–FGFR3c structure.

To confirm that the additional interaction between the N terminus of FGF1 and FGFR3c observed in the crystal structure also occurs in solution, we compared the kinetics of the interaction between full-length and truncated FGF1 with the three-Ig form of FGFR3c by using SPR. Importantly, whereas truncated FGF1 negligibly binds to the three-Ig form of FGFR3c, full-length FGF1 binds with a dissociation constant (K_d) of 916 nM (Fig. 2 *A* and *B*). Similar results were obtained with the two-Ig form of FGFR3c, namely, truncated FGF1 binds FGFR3c with a 3,000-fold weaker affinity compared with full-length FGF1 (Table 2). This finding is consistent with our inability to isolate a stable complex of truncated FGF1 with either the two- or three-Ig form of FGFR3c on a size-exclusion column. Taken together, the data suggests that the N-terminal interactions observed in the crystal structure also occur in solution.

Next, we assessed the biological relevance of the contacts between N-terminal residues of FGF1 and FGFR3c. To do so, we compared the ability of full-length and truncated FGF1 to induce FGFR3c autophosphorylation and extracellular signal-regulated kinase (ERK)1/2 activation in RCS cells. Consistent with the SPR and structural data presented above, full-length FGF1 was more potent than truncated FGF1 in activating FGFR3 and ERK1/2 relative to truncated FGF1 (Fig. 2 *C* and *E*).

Due to the fact that truncated FGF1 (residues 21–155) was used in the FGF1–FGFR1c and FGF1–FGFR2c structures, it is not clear whether additional interactions involving the FGF1 N terminus are also important in FGFR2 binding and activation. Because RCS cells also express FGFR2, we analyzed the phosphorylation of FGFR2 in response to full-length and truncated FGF1. Interestingly, full-length FGF1 was also more potent than truncated FGF1 in stimulating FGFR2 phosphorylation (Fig. 2*D*). Therefore, it is likely that additional N-terminal residues in FGF1 that are not included in the truncated construct, also participate in FGFR2 binding and activation.

D1 and the D1–D2 Linker Negatively Regulate the Binding Affinity of FGFR3c for Both FGF1 and Heparin. Having ruled out the possibility that D1 or the D1–D2 linker region play a positive role in FGF1 binding, it was necessary to address the proposed negative regulatory role of D1 and the D1–D2 linker in FGFR signaling (13, 18). We used SPR to compare the kinetics of the interaction between the two-Ig and the three-Ig domain form of FGFR3c with full-length FGF1 and heparin. For studies of the FGF1–FGFR3c interaction, varying concentrations of FGFR were injected over a protein sensor chip onto which FGF1 was immobilized, and for studies of the FGFR–heparin interaction, various concentrations of FGFR were injected over a heparin neoproteoglycan sensor chip. As shown in Fig. 3 *A* and *B*, the three-Ig form of FGFR3c binds to FGF1 with an ≈4-fold lower affinity than did the two-Ig form. The

three-Ig form of FGFR3c also bound with a 5.7-fold lower affinity for heparin than did the two-Ig form (Fig. 3 *C* and *D*). The reduced binding affinity of the three-Ig form of FGFR3c for FGF1 and heparin is primarily due to a decrease in the association rate (k_{on}) of the interaction (Table 2).

Our findings are qualitatively consistent with the results of Scatchard analysis of FGF1 and heparin binding to mammalian and insect cells expressing the three- or two-Ig form of FGFR1c. It was shown that the three-Ig form of FGFR1 has a 6.7- and 2.6-fold lower affinity for FGF1 and heparin, respectively, relative to the two-Ig form (18). However, the calculated affinities for the interaction of three- and two-Ig forms of FGFR1 with FGF1 in this Scatchard analysis study are roughly two orders of magnitude higher than those determined with SPR. The elevated binding affinity of FGF1 toward FGFR in this Scatchard analysis study is probably due to difference in the state of receptor (cell surface-bound versus in solution) and/or the presence of soluble heparin in the binding experiments, which is known to enhance the binding affinity of FGF for FGFR. It is also important to note that the calculated affinities obtained with SPR in this study are quantitatively similar to those obtained with isothermal titration calorimetry of the interaction between soluble FGFR and FGF in the absence of heparin (26).

Mechanism of FGFR Autoinhibition. By using a panel of Ig domain-specific monoclonal and polyclonal antibodies against FGFR1, Mckeehan and coworkers (18, 27) have proposed that D1 interacts with the D2–D3 fragment of FGFR, thereby inhibiting FGF and heparin binding. We used SPR to determine whether a construct encompassing D1 and the D1–D2 linker of FGFR3c (D1+) interacts directly with the D2–D3 fragment of FGFR3c. To that end, varying concentrations of D1+ were injected over a protein sensor chip onto which FGFR3c was immobilized. D1+ was found to interact with the two-Ig form of FGFR3c with a dissociation constant (K_d) of $\approx$20 μM (Fig. 3*E*).

Cumulatively, the above data substantiate a model in which D1 and the D1–D2 linker negatively regulate FGFR activation. The flexible nature of the D1–D2 linker facilitates a pair of intramolecular interactions that include the acid box engaging in electrostatic interactions with the highly basic heparin-binding site of D2, and D1 interacting directly with the D2–D3 fragment of FGFR. These interactions keep FGFR in a "closed" low-activity state, in which the affinity of FGFR for both FGF and heparin are reduced. This model is consistent with the finding that both the reduction in the affinity of the three-Ig form of FGFR3c for heparin and FGF1 are primarily due to a decrease in the association kinetics of the interaction. We suggest that the interactions of the acid box with the heparin-binding site, and D1 with the D2–D2 fragment of FGFR are individually weak, but synergize to sustain receptor autoinhibition. Importantly, this model for FGFR autoinhibition is consistent with the disordering of D1 and the D1–D2 linker in the FGF1–FGFR3c structure, because these autoinhibitory interactions must be disrupted in order for ligand binding to take place.

It is interesting to note that D1 and the D1–D2 linker region are the least conserved extracellular subdomains of FGFRs. Moreover, both the length of the D1–D2 linker and the number of the consecutive acidic residues, which define each acid box, differ significantly among the four FGFRs (3). Therefore, we propose that gradations of autoinhibitory control exist for each FGFR. FGFR1, which has both the longest D1–D2 linker and acid box, is likely to be under strong autoinhibition. In contrast, FGFR4, which has the shortest acid box, is probably under mild autoinhibition. Analysis of tissue distribution of components of FGF signal transduction provides plausible explanation for why FGFR activity needs to be under such tight autoinhibitory control. Unlike many signaling systems, components of FGF signaling are often colocalized within tissues. Heparan sulfate proteoglycans are ubiquitously expressed and bind FGFs with high affinity and capacity, thereby sequestering them in the vicinity of FGFRs. Thus, the probability of FGF- or heparin-independent dimerization or inadvertent ternary complex formation and dimerization is extremely high, and necessitates tight autoinhibitory control.

We thank Drs. C. Ogata, X. Yang, and R. Abramowitz for synchrotron beamline assistance; Y. Lu (New York University Protein Analysis Facility) for mass spectrometric analysis; Drs. C. Boesen and X. Kong for providing useful comments on the manuscript; and Dr. B. Yeh and N. Takeda for assisting with figure preparation. Beamline X4A at the National Synchrotron Light Source, a U.S. Department of Energy facility, is supported by the Howard Hughes Medical Institute. This work was supported by National Institutes of Health Grants DE13686 (to M.M.), DE013745 (to C.B.), and HL52622 (to R.J.L.).

1. Ornitz, D. M. & Itoh, N. (2001) *Genome Biol.* **2,** 1–12.
2. Givol, D. & Yayon, A. (1992) *FASEB J.* **6,** 3362–3369.
3. Johnson, D. E. & Williams, L. T. (1993) *Adv. Cancer Res.* **60,** 1–41.
4. McKeehan, W. L., Wang, F. & Kan, M. (1998) *Prog. Nucleic Acid Res. Mol. Biol.* **59,** 135–176.
5. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P. & Ornitz, D. M. (1991) *Cell* **64,** 841–848.
6. Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E. & Leder, P. (1992) *Mol. Cell. Biol.* **12,** 240–247.
7. Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J. & Mohammadi, M. (2000) *Mol. Cell* **6,** 743–750.
8. Miki, T., Bottaro, D. P., Fleming, T. P., Smith, C. L., Burgess, W. H., Chan, A. M. & Aaronson, S. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 246–250.
9. Yayon, A., Zimmer, Y., Shen, G. H., Avivi, A., Yarden, Y. & Givol, D. (1992) *EMBO J.* **11,** 1885–1890.
10. Dell, K. R. & Williams, L. T. (1992) *J. Biol. Chem.* **267,** 21225–21229.
11. Ornitz, D. M., Xu, J., Colvin, J. S., McEwen, D. G., MacArthur, C. A., Coulier, F., Gao, G. & Goldfarb, M. (1996) *J. Biol. Chem.* **271,** 15292–15297.
12. Shimizu, A., Tada, K., Shukunami, C., Hiraki, Y., Kurokawa, T., Magane, N. & Kurokawa-Seo, M. (2001) *J. Biol. Chem.* **276,** 11031–11040.
13. Plotnikov, A. N., Schlessinger, J., Hubbard, S. R. & Mohammadi, M. (1999) *Cell* **98,** 641–650.
14. Plotnikov, A. N., Hubbard, S. R., Schlessinger, J. & Mohammadi, M. (2000) *Cell* **101,** 413–424.
15. Stauber, D. J., DiGabriele, A. D. & Hendrickson, W. A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 49–54.
16. Yeh, B. K., Igarashi, M., Eliseenkova, A. V., Plotnikov, A. N., Sher, I., Ron, D., Aaronson, S. A. & Mohammadi, M. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 2266–2271.
17. Chellaiah, A., Yuan, W., Chellaiah, M. & Ornitz, D. M. (1999) *J. Biol. Chem.* **274,** 34785–34794.
18. Wang, F., Kan, M., Yan, G., Xu, J. & McKeehan, W. L. (1995) *J. Biol. Chem.* **270,** 10231–10235.
19. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
20. Navaza, J. (1994) *Acta Crystallogr. A* **50,** 157–163.
21. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905–921.
22. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, G. (1991) *Acta Crystallogr. A* **47,** 110–119.
23. Ibrahimi, O. A., Zhang, F., Eliseenkova, A. V., Linhardt, R. J. & Mohammadi, M. (2004) *Hum. Mol. Genet.* **13,** 69–78.
24. Raucci, A., Laplantine, E., Mansukhani, A. & Basilico, C. (October 30, 2003) *J. Biol. Chem.* 10.1074/jbc.M310384200.
25. Eriksson, A. E., Cousens, L. S., Weaver, L. H. & Matthews, B. W. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3441–3445.
26. Pantoliano, M. W., Horlick, R. A., Springer, B. A., Van Dyk, D. E., Tobery, T., Wetmore, D. R., Lear, J. D., Nahapetian, A. T., Bradley, J. D. & Sisk, W. P. (1994) *Biochemistry* **33,** 10229–10248.
27. Xu, J., Nakahara, M., Crabb, J. W., Shi, E., Matuo, Y., Fraser, M., Kan, M., Hou, J. & McKeehan, W. L. (1992) *J. Biol. Chem.* **267,** 17792–17803.
28. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
29. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).