Crystal Structure of Fibroblast Growth Factor 9 Reveals Regions Implicated in Dimerization and Autoinhibition*

Received for publication, July 20, 2000, and in revised form, October 30, 2000
Published, JBC Papers in Press, November 1, 2000, DOI 10.1074/jbc.M006502200

Alexander N. Plotnikov‡, Anna V. Eliseenkova‡, Omar A. Ibraimi‡, Zachary Shriver§,
Ram Sasisekharan§, Mark A. Lemmon¶, and Moosa Mohammadi∥

From the ‡Department of Pharmacology, New York University School of Medicine, New York, New York 10016, the §Department of Biochemistry and Biophysics and the Johnson Foundation, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059, and the ¶Division of Bioengineering and Environmental Health, Harvard-Massachusetts Institute of Technology, Division of Health Sciences and Technology, Cambridge, Massachusetts 02139

Fibroblast growth factors (FGFs) constitute a large family of heparin-binding growth factors with diverse biological activities. FGF9 was originally described as a glia-activating factor and is expressed in the nervous system as a potent mitogen for glia cells. Unlike most FGFs, FGF9 forms dimers in solution with a $K_d$ of 680 nM. To elucidate the molecular mechanism of FGF9 dimerization, the crystal structure of FGF9 was determined at 2.2 Å resolution. FGF9 adopts a β-trefoil fold similar to other FGFs. However, unlike other FGFs, the N- and C-terminal regions outside the β-trefoil core in FGF9 are ordered and involved in the formation of a 2-fold crystallographic dimer. A significant surface area (>2000 Å²) is buried in the dimer interface that occludes a major receptor binding site of FGF9. Thus, we propose an autoinhibitory mechanism for FGF9 that is dependent on sequences outside of the β-trefoil core. Moreover, a model is presented providing a molecular basis for the preferential affinity of FGF9 toward FGFR3.

The mammalian fibroblast growth factor (FGF) family contains at least 22 distinct polypeptides (FGF1–FGF22) that are expressed in a specific spatial and temporal pattern (1–5). FGFs play important roles in numerous physiological and pathological processes (1–5). Members of the FGF family share between 10 and 55% sequence identity (6). Crystal structures of FGF1 (7), FGF2 (7–9), and FGF7 (10), have revealed a common core FGF structure consisting of three copies of a four-stranded β-sheet, known as a β-trefoil fold. All FGFs contain N- and C-terminal segments outside the β-trefoil core. However, the length of these segments, especially the C termini, varies greatly among different FGFs. In FGF1, FGF2, FGF4, and FGF7, the end of polypeptide chain virtually coincides with the end of the β-trefoil fold. Consequently, these FGFs have extremely short C-terminal segments. In contrast, other FGFs, such as FGF3, FGF5, and FGF8, have long C-terminal extensions. The functional relevance of these N- and C-terminal extensions remains uncharacterized.

FGF9 (glia-activating factor) was purified as a heparin-binding, secreted glycoprotein from cultured human glioma cell line NMC-G1 (11). FGF9 shows 30% sequence identity with the prototypical FGF family members, FGF1 and FGF2 (12). Purified FGF9 is mitogenic for many types of cultured cells including glia cells, oligodendrocyte type 2 astrocyte progenitor cells, smooth muscle cells, pheochromocytoma PC12 cells, and BALB/3T3 fibroblasts (11). Because FGF9, unlike FGF1 and FGF2, has no effect on human umbilical vein endothelial cells, it is suggested that FGF9 may have a unique receptor specificity (11). In fact, biochemical studies performed utilizing various soluble FGFR-alkaline phosphatase fusion proteins and genetically engineered cells expressing different full-length FGFs have demonstrated that FGF9 binds preferentially to the IIIc form of FGFR3 (13–15).

FGF9, like prototypical FGFs, does not have a typical secretory signal peptide. Yet, it is still glycosylated and efficiently secreted from transfected mammalian cells (12). N-terminal sequencing of secreted FGF9 shows that it is missing only the initiation methionine (12, 16). It has been suggested that FGF9 secretion is mediated via a noncleaved signal sequence consisting of the first 33 residues (17, 18). We have expressed and purified truncated FGF9 (residues 35–208) in Escherichia coli and observed that it dimerizes in solution. To elucidate the molecular mechanism of FGF9 dimerization, the crystal structure of FGF9 was determined. In the crystal, FGF9 forms a 2-fold dimer in which a large surface area is buried. This suggests that the dimer observed in the crystal structure also represents the dimeric FGF9 species detected in solution. Analysis of the dimer interface reveals interactions between the N- and C-terminal segments outside of the β-trefoil core of each FGF9 protomer that are the driving force for dimer formation. Interestingly, a major FGFR binding site becomes occluded upon dimer formation. Thus, we propose that FGF9 regions outside the β-trefoil core have an autoinhibitory role in FGF9 function. Furthermore, our modeling studies afford a structural basis for the preferential affinity of FGF9 toward FGFR3 and for its specific heparin requirement.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—DNA fragments generated by polymerase chain reaction of human FGF9 cDNA (encoding for residues 35–208) were subcloned into the pET-28a bacterial expression vector using Ncol and HindIII cloning sites. Following transformation of the BL21 (DE3) bacterial strain, cells containing the FGF9-expression plasmid were induced with 1 mM isopropyl-1-thio-D-galactopyranoside for 5 h. The bacteria were then centrifuged and subsequently lysed in a 25 mM HEPES-NaOH buffer (pH 7.5) containing 150 mM NaCl, 10% glycerol, and 5 mM EDTA using a French press. FGF9 was found mainly in the soluble fraction and precipitated with saturated ammonium sulfate.
solution (final concentration, 50% overnight at 4 °C. Following centrifugation, the pellet was dissolved in a HEPES-NaOH buffer (pH 7.5) containing 1 M ammonium sulfate and diluted 1:5 prior to loading onto a Source S column (Amersham Pharmacia Biotech). Bound FGF9 was eluted by a linear gradient of NaCl to 1 M in a 25 mM HEPES-NaOH buffer (pH 7.5) buffer. Subsequent purification of FGF9 was achieved by size exclusion chromatography on a Superdex 75 column (Amersham Pharmacia Biotech) equilibrated with a 25 mM HEPES-NaOH buffer (pH 7.5) containing 1.5 M NaCl. Mass spectrometry of purified FGF9 confirmed the predicted molecular mass of 20,169 daltons for FGF9.

Crystal Structure of FGF9—Crystals of FGF9 were grown by vapor diffusion at 20 °C using the hanging drop method. 2 mM protein solution (1.4 mg/ml in 25 mM HEPES-NaOH, pH 7.5, and 150 mM NaCl) were mixed with 8 μl of various crystallization buffer conditions of Crystal Screen kits (Hampton Research). Condition 48 (kit 1) (2 mM ammonium phosphate, 0.1 M Tris-HCl, pH 8.5) produced diffraction quality crystals. FGF9 crystals belong to the tetragonal space group I4122 with unit cell dimensions a = b = 87.82 Å, c = 115.64 Å. There is one molecule of FGF9 in the asymmetric unit with a solvent content of ~56%. Diffraction data were collected from a flash frozen (in a dry nitrogen stream using mother liquor containing 20% glycerol as cryoprotectant) crystal on a CCD detector at beamline X4A at the National Synchrotron Light Source, Brookhaven National Laboratory. Data were processed using ENZOE and SCALEPACK (19).

Structure Determination and Refinement—A molecular replacement solution was found for one copy of FGF9 in the asymmetric unit using the program AmoRe (20) and the structure of FGF4 as the search model. Simulated annealing and positional/B factor refinement were performed using CNS (21). Bulk solvent and anisotropic B factor corrections were applied. Model building into 2Fo – Fc and Fc – Fo electron density maps was performed with program O (22). The atomic model contains residues 52–208 of FGF9, two phosphate ions, and 68 water molecules. The average B factor is 18.5 Å² for FGF9 molecule, 56.0 Å² for phosphate ions, and 24.0 Å² for water molecules.

Matrix-assisted Laser Desorption Ionization Mass Spectrometry Analysis of FGF9—Mass spectral analysis of FGF9 was completed by diluting 1:10 of an aqueous 10 mM stock solution of FGF9 with a saturated solution of sinapinic acid in 30% acetonitrile. The sample was spotted on a seeded surface and dried under a stream of nitrogen (23). MALDI-MS were acquired in the linear mode using delayed extraction with a Voyager Elite reflection time-of-flight instrument (PerSeptive Biosystems, Framingham, MA) with a 337-nm laser. Spectra were acquired using instrument settings described previously (24). Mass spectra were calibrated externally with myoglobin and bovine serum albumin.

Sedimentation Equilibrium Ultracentrifugation Analysis—Sedimentation equilibrium experiments employed an XL-A analytical ultracentrifuge (Beckman). Samples (25 mM HEPES, pH 7.5, containing 1.5 M NaCl) were loaded into six-channel epon charcoal-filled centerpieces, using quartz windows. Experiments were performed at 20 °C, detecting at 235 and 278 nm, using four different speeds (12,000, 18,000, 24,000, and 30,000 rpm.). The program SEDNTERP was used to estimate the volume of FGF9 (0.732 ml/g) from its amino acid composition (25). Experiments were performed with a range of FGF9 concentrations ranging from 1.6 to 24 μM at the speeds mentioned above. The Optima XL-A Data Analysis Software package (Beckman/Microcal) was used for global fits of the data to self-association models using nine data sets with 278 nm detection and (independently) nine data sets with 235 nm detection. Identical results were obtained for each group of nine data sets. Reasonable global fits to the data could be obtained only with a simple dimerization model, as discussed under “Results and Discussion.” Goodness of fit was judged by the occurrence of small, randomly distributed, residuals for the fits, as shown in Fig. 2B.

Results and Discussion

Structure Determination—The first 33 N-terminal residues of FGF9 were deleted because they are implicated only in secretion and not FGFR binding (17, 26). Truncated FGF9 (residues 35–208) was expressed in E. coli and purified to homogeneity (see “Experimental Procedures”). Crystallization trials with FGF9 produced tetragonal crystals with one molecule per asymmetric unit. The crystal structure of FGF9 was solved by molecular replacement (see “Experimental Procedures”) and refined at 2.2 Å resolution with an R value of 20.7% (free R value of 23.1%). The atomic model for FGF9 consists of one FGF9 molecule (residues 52–208), two phosphate ions, and 68 water molecules. Data collection and refinement statistics are given in Table I.

Description of the Structure—As anticipated on the basis of sequence alignment, FGF9 adopts a β-trefoil fold (Fig. 1). Superimposition of the Cα traces located within the β-trefoil core of FGF9 with those belonging to FGF1 and FGF2 gives a root-mean-square deviation of only 1.0 and 0.73 Å, respectively. Within the β-trefoil core, the major differences between the FGF9 structure and the structures of FGF1 and FGF2 are the β1-β2 and β9-β10 loop conformations. These loops vary both in

![Ribbon diagram of FGF9](image-url)
length and sequence among the various members of the FGF family. In FGF9, the β1-β2 loop is one residue shorter than the corresponding loops of FGF1 and FGF2. In contrast, the β9-β10 loop in FGF9 is longer by 4 and 6 residues than the corresponding loop in FGF1 and FGF2, respectively. This loop bulges out from the main β-trefoil body of FGF9. Two ordered phosphate ions are coordinated in the FGF9 structure; one is bound to the high affinity heparin-binding site at a position similar to where sulfate ions/groups bind ligand in the structures of free or heparin-bound FGF1 (27, 28) and FGF2 (8, 9, 29). The other
phosphate ion is bound by FGF9 residues, whose homologues in FGF1 and FGF2 have not been implicated in sulfate ion coordination. The most striking structural differences between FGF9 and the prototypical FGFs are the conformations of the N- and C-terminal regions outside of the β-trefoil core. In the crystal structures of free FGF1 and FGF2, the region preceding the β-trefoil core is disordered. However, in the FGF9 structure, part of this region is ordered and forms an α-helix (αN) (Fig. 1). Additionally, the entire C-terminal tail following the β-trefoil core is also ordered and contains a short helix (αC) (Fig. 1). The C-terminal region is considerably longer in FGF9 (17 residues) than the C-terminal regions of FGF1 (4 residues), and FGF2 (3 residues). Several interhelical interactions between the αN and αC helices contribute to the ordering of the N- and C-terminal segments in FGF9.

**FGF9 Dimers and Autoinhibition**—As shown in Fig. 2A, the elution position of FGF9 on the size exclusion chromatography column is dependent on the concentration at which it is loaded. When loaded at 2 mg/ml (100 μM), FGF9 elutes at a position corresponding to a molecular mass of ~40 kDa, indicating that FGF9 forms dimers. By contrast, when loaded at a concentration of 8 μg/ml (0.4 μM), FGF9 elutes at 32 ml, consistent with its monomeric molecular mass (20 kDa). When concentrations between 5 μg/ml and 2 mg/ml are loaded, the elution position of FGF9 is intermediate, suggesting that FGF9 reversibly dimerizes with a $K_d$ in the micromolar range. This was confirmed quantitatively by sedimentation equilibrium analytical ultracentrifugation studies (Fig. 2B), which showed that FGF9 dimerizes with a $K_d$ of 680 nM. Sedimentation of FGF9 was analyzed for a series of solutions at concentrations from 1.6 to 26 μM, as described under “Experimental Procedures.” Attempts to fit the data to a single species suggested that FGF9 self-associates, and only a simple dimerization model could be fitted globally to multiple data sets collected at different concentrations and speeds. The global fit to the data shown in Fig. 2B represents a model in which FGF9 dimerizes with a $K_d$ of 680 nM. As seen above each graph, the residuals for this fit are small and random for each individual data set.

Additional verification of the ability of FGF9 to dimerize was achieved using MALDI-MS under conditions that do not interfere with protein-protein interactions (30, 31). Using sinapinic acid as the matrix, two FGF9 species are detected, with molecular masses of 20 and 40 kDa (Fig. 2C). Use of other matrix conditions, including ferulic acid in tetrahydrofuran and 6-aza-2-thiothymine in ammonium acetate, confirms that the dimer observed in the mass spectrum is not simply an artifact of MALDI-MS sample preparation (data not shown) and provides further support for FGF9 dimerization in solution.

Information regarding possible structural modes of FGF9 dimerization can be gained by close inspection of the FGF9 crystal structure. Analysis of the symmetry mates in the crystal reveals a 2-fold crystallographic dimer that is particularly intimate (Fig. 3). The large buried surface area (2260 Å²) in the interface between molecules in this dimer is compatible with our measured dissociation constant of 680 nM, suggesting that this is the dimer that occurs in solution.

The dimer interface can be analyzed as the sum of two interfaces: an interface outside the β-trefoil core (Fig. 4A) and an interface inside the β-trefoil core (Fig. 4B). The interface outside the β-trefoil involves the αN and αC helices as well as the respective loops that connect them to the main β-trefoil body. At this interface, the aliphatic side chains of Leu-54, Ile-60, Leu-61, Leu-200, and Ile-204 from one FGF9 protomer are in hydrophobic contact with the corresponding residues of the other protomer (Fig. 4A). A number of hydrogen bonds further strengthen this interface. Sequence alignment of FGFs shows that most of the residues that participate in this interface are conserved in FGF16, suggesting that FGF16 may also dimerize in solution (Fig. 5). Analogous residues in the remain-
ing FGFs are divergent, implying that dimerization of these ligands by a similar mechanism is not likely.

The portion of the dimer interface inside the β-trefoil core involves the β1 and β12 strands as well as the β8-β9 loop (Fig. 4B). Several hydrogen bonds fortify hydrophobic contacts that are made at this interface (Fig. 4B). Surprisingly, several FGF9 residues that participate in this interface are also predicted to interact with Ig-like domain 2 (D2) of FGFR in an FGF9-FGFR complex. As expected, these residues are highly conserved among different FGFs. Still, solution dimers for other FGFs have not been reported, insinuating that these interactions alone are not sufficient to drive FGF dimerization. Therefore, interactions between FGF9 protomers at the dimer interface outside the β-trefoil core are essential in promoting FGF9 dimerization. These interactions must facilitate the formation of the dimer interface within the β-trefoil core and cooperate with this second interface for dimerization to occur. In the process, critical receptor binding sites appear to become occluded upon dimerization, suggesting that dimerization may serve as a biologically relevant autoinhibitory mechanism.

Receptor Binding Site and Specificity—Receptor binding specificity is an essential regulatory element of FGF responses and is achieved through both primary sequence variations and alternative splicing. Comparison of three different FGF-FGFR structures defined a general binding interface for FGF-FGFR complexes that involves contacts made by FGF to D2 and to the linker between D2 and the Ig-like domain 3 (D3) (32). It was also shown that specificity is achieved through interactions between the FGF N-terminal (immediately preceding the β-trefoil core) and central regions with two loop regions in FGFR D3 that are subject to alternative splicing (32).

To provide a molecular basis for the specificity of FGF9 toward FGFR3(IIIc), we superimposed the FGF9 structure onto the FGF2 structure complexed with the ligand binding portion of FGFR1 consisting of D2 and D3 (Fig. 6A). The majority of the interactions in the FGF-FGFR interface can easily be accommodated with minor adjustments of side chain rotamers. However, analysis of the FGF9-FGFR1 interface identified two critical regions, the β8-β9 loop (within the β-trefoil core) and the αN helix (at the N terminus outside the β-trefoil core), that are sterically clashing with the receptor (colored red in Fig. 6A). These two regions of FGF9 would require major changes in
backbone conformation to allow an engagement with FGFR1 to occur.

At the primarily hydrophobic FGF9-D2 interface, three critical FGF9 residues (Tyr-67, Tyr-145, and Leu-188) are highly conserved among different members of the FGF family (Fig. 6). In the FGF9-linker interface, Asn-146, also highly conserved among FGFs, is expected to form hydrogen bonds with an FGFR-invariant arginine (Arg-250 in FGFR1) in the D2-D3 interface.

**FIG. 5.** Structure-based sequence alignment of FGFs. Sequence alignment was performed using the CLUSTALW (41). All of the FGFs used in this alignment are human. The location and the length of the β strands and α helices are shown on the top of the sequence alignment. A period indicates sequence identity to FGF9. A dash represents a gap introduced to optimize the alignment. FGF9 residues that participate in dimerization are colored red. In blue are FGF9 residues that constitute the conventional low and high affinity heparin binding sites. FGF9 residues that localize to the periphery of the high affinity heparin-binding site and are predicted to form the distal and central heparin binding sites are colored green and yellow, respectively.

**FIG. 6.** Mapping of receptor binding sites in FGF9. A, a model for the FGF9-FGFR1 structure was generated by superimposition of the Cα traces within the β-trefoil of the FGF9 structure onto the corresponding Cα traces of FGF2 in the FGF2-FGFR1 structure. Orange, FGF9; green, D2; cyan, D3; gray, linker region; red, FGF9 regions that are in major clashes with FGFR1. B, stereo view of the receptor binding sites on FGF9. FGF9 residues are colored with respect to the FGFR regions with which they interact. FGF9 residues that interact with D2 are colored green, residues that interact with the linker region are colored gray, and residues that interact with D3 are colored cyan. FGF9 residues that interact with the βC9βE loop in D3 of FGFR are colored purple. Color coding for atoms is the same as Fig. 4. This figure was created using Molscript and Raster3D.
linker region. In contrast, FGF9 residues predicted to be in the FGF9-D3 interface show little conservation among FGFs. Close inspection of the FGF9-D3 interface at the $\beta C$-$\beta E$ loop region in D3 affords a potential explanation for why FGF9 binds preferentially to FGFR3(IIIc) over FGFR1(IIIc). The $\beta C$-$\beta E$ loops of the two FGFRs differ at two positions. Significantly, the residue corresponding to Val-316 in FGFR1 is an alanine (Ala-313) in FGFR3. In the FGF2-FGFR1 structure, Val-316 makes hydrophobic contacts with Tyr-73, Val-88, and Phe-93 in FGF2. Although Tyr-73 is conserved in all FGFs including FGF9, Leu-130 and Val-135 in FGF9 replace Val-88 and Phe-93 in FGF2 (Fig. 6B). The larger hydrophobic side chain of Leu-130 in FGF9 clashes sterically with Val-316 in FGFR1, thus reducing the affinity of FGF9 toward FGFR1(IIIc). Conversely, the bulkier Leu-130 side chain would better engage the smaller side chain of Ala-313 in the FGF9-FGFR3 model. Interestingly, two other FGF9 residues (Ile-98 and Ile-100) also fall in the vicinity of Ala-313 in the FGF9-FGFR3 model. The aliphatic side chains of these two residues may provide additional hydrophobic contacts in both FGF2-FGFR1 and FGF2-FGFR2 complexes (32, 33).

Heparin Binding Site—The requirement of heparin/heparan sulfate proteoglycans for FGF-induced FGFR dimerization is well documented (34). According to the proposed “two-end” model, heparin interacts via its nonreducing end with the heparin binding sites of FGF and FGFR and promotes the formation of a ternary 1:1:1 FGF-FGFR-heparin complex. A second ternary 1:1:1 FGF-FGFR-heparin complex is then recruited to the first ternary complex via interactions of FGFR, FGF, and heparin in one ternary complex with the FGFR in the adjoining ternary complex.

Within each ternary 1:1:1 FGF-FGFR-heparin complex, heparin makes numerous contacts with the heparin binding residues of FGF and FGFR, thereby increasing the affinity of FGF toward FGFRs. This provides the molecular basis for the well-documented heparin-dependent 1:1 FGF-FGFR interaction. In addition, heparin also interacts with the heparin binding site of the adjoining FGFR, thereby augmenting the weak interactions of FGF and FGFR in one ternary complex with the FGFR in the adjoining ternary complex. In the absence of heparin, these interactions are not sufficient for sustained dimerization to occur. Based on the model, differences in the primary sequences of heparin binding sites of FGFs and FGFRs will effect heparin-dependent FGF-FGFR binding affinity. Hence, each FGF may require different heparin motifs (sulfation pattern and/or length) to exert their optimal biological activities.

To analyze potential heparin binding sites of FGF9, a dimeric model for FGF9-FGFR1 was generated by superimposing two copies of the FGF9 structure onto the two copies of FGF2 in the ternary FGF2-FGFR1-heparin structure. As in the FGF2-
FGFR1 structure, calculation of the electrostatic surface potential identifies a positively charged canyon. Two phosphate ions (provided by the crystallization buffer) are bound in the predicted high affinity heparin binding sites of FGF9 (Fig. 7A). Because heparin or heparan sulfate moieties of cell surface proteoglycans are certainly longer than heparin decasaccharides used in the FGFR2-FGFR1-heparin crystals (33), it is reasonable to suggest that natural heparins may interact with FGFs beyond the well characterized low and high affinity heparin binding sites. Indeed, close inspection of surface residues at the periphery of the conventional FGF high affinity site led us to consider two unreported potential sites: one at the very distal end of the canyon and another that is more central to high affinity site. The distal site, composed of Thr-81, Lys-121, and Glu-123, does not coordinate a phosphate ion in the FGF9 structure (Fig. 7B). However, the crystal structure of FGF4 reveals a sulfate ion bound at this site by side chains of the corresponding FGF4 residues (Arg-103, Lys-142, and Lys-144). Thompson et al. (36) who mutated several surface residues of FGF2 to identify potential heparin binding sites, have provided further evidence for the existence of a distal site. Mutation of Arg-81 in FGF2, which corresponds to Glu-123 in FGF9, was found to cause a 2-fold reduction in heparin binding.

In contrast, the “central” site does coordinate a phosphate ion via the side chains of Tyr-163 and Arg-180 (Fig. 7B). Several other FGF9 residues in the vicinity of this phosphate ion may also interact with heparin (Fig. 7B). Interestingly, several of these residues, including Tyr-163, are located in the highly divergent β9-β10 loop. Because bound sulfate ions at the high affinity heparin binding site of FGFs often mimic the position of the sulfate groups of heparin in heparin-bound FGFs, it is reasonable to suggest that this central site may indeed constitute a genuine heparin binding site. Confirmation of the aforementioned hypothesis requires structural determination of a ternary FGF9-FGFR3-heparin complex.

Given the importance of FGF signaling in numerous biological processes, it is likely that multiple levels of regulation exist to modulate FGF activity. Besides FGF9 and FGF16, several other uncharacterized members of the FGF family such as FGF17, FGF18, and FGF19 contain large N- and C-terminal regions outside the β-trefoil main body. It is possible that these regions may also possess some type of autoregulatory activity. Structural delineation of such regions in FGFs may be particularly significant for the design of FGF mimetics. FGF agonists may be used therapeutically to accelerate wound healing or to induce angiogenesis in pathological conditions such as angina and stroke.