X-ray Crystal Structure of Human Acidic Fibroblast Growth Factor[†,‡]

Michael Blaber,*[,§] Jerry DiSalvo,[||] and Kenneth A. Thomas[||]

Institute of Molecular Biophysics and Department of Chemistry, Florida State University, Tallahassee, Florida 32306-3015, and Department of Biochemistry, Merck Research Laboratories, P.O. Box 2000, Rahway, New Jersey 07065

Received September 12, 1995; Revised Manuscript Received November 29, 1995[⊗]

ABSTRACT: Fibroblast growth factors (FGFs) are mitogenic and chemotactic agents for a wide variety of cell types and play a primary role in the regulation of angiogenesis. Angiogenesis is involved in a variety of critical physiological events including organogenesis, wound healing, ischemic collateral circulation, and solid tumor growth. High-resolution structural information is key to understanding the mechanism of action of these growth factors. We report here the X-ray crystal structure of human acidic FGF (aFGF), with data extending to 2.0 Å resolution. The crystal contains four independent molecules in the asymmetric unit. Each molecule contains a single bound sulfate ion, in similar juxtapositions. The bound sulfate is stabilized through hydrogen-bond interactions with residues Asn 18, Lys 113, and Lys 118 and defines a potential heparin binding site. The hydrogen bond with the Nδ_2 moiety of Asn 18 appears to be the most conserved interaction, being similar to those observed for sulfate ion bound to human basic FGF (bFGF) and similar but not identical to interactions observed for bovine aFGF with heparin analogs. Of the added solvent groups, five ordered water molecules are conserved in each of the four independent structures of human aFGF. These water molecules, located at buried positions, provide hydrogen bonding partnerships with several buried polar groups in the core of the protein. A central interior cavity exists in each of the four structures, with sizes ranging from approximately 20 to 50 Å^3. The cavity sizes appear to be significantly smaller than that observed in the related protein interleukin-1β. The region comprising the high affinity FGF receptor binding site is structurally very similar to the corresponding region from human bFGF, whereas the low affinity site is structurally quite different. The results provide a structural basis for the role of the low affinity binding site in FGF receptor discrimination.

Acidic FGF[1] is one of the more extensively characterized of the nine known FGFs (Burgess & Maciag, 1989; Miyamoto et al., 1993; Tanaka et al., 1992; Thomas, 1993). It is the only FGF known to bind with high affinity to all four identified FGF membrane-spanning tyrosine kinase receptors (FGFRs) and their variants, which are generated by alternative mRNA splicing and exhibit altered ligand-binding selectivities (Chellaiah et al., 1994). The wide distribution of FGFR expression among various cell types, including most, if not all, cells of mesodermal and ectodermal embryonic origin, coupled with the ability of aFGF to tightly bind and presumably function through all known tyrosine kinase FGF receptor forms, indicates that it is possibly the broadest spectrum growth factor yet characterized.

As with the other characterized FGFs, aFGF binds to heparin and heparan proteoglycans. Although the biological significance of binding to these polyanions is not yet clear, heparin binding is thought to be required for binding of bFGF to its high affinity receptor by formation of a ternary complex among ligand, receptor, and heparin (Ornitz et al., 1992; Rapraeger et al., 1991; Yayon et al., 1991). However, recent evidence suggests that heparin might not be required for binding but, perhaps, stabilizes either the conformation of bFGF that has been oxidatively damaged by the particular method used to iodinate the protein (Roghani et al., 1994) or the ligand−receptor complex (Pantoliano et al., 1994). Binding to heparin clearly stabilizes aFGF thereby diminishing inactivation by heat, acid (Gospodarowicz & Cheng, 1986), proteolysis (Rosengart et al., 1988), and mild oxidation *in vitro* (Linemeyer et al., 1990). Stabilization is particularly relevant for human aFGF, which exhibits an unfolding transition at physiological temperatures in the absence of these polyanions (Copeland et al., 1991). In contrast to the highly homologous bovine protein (Linemeyer et al., 1987), the mitogenic activity of human aFGF in culture is also remarkably dependent on exogenous heparin (Ortega et al., 1991). One physiological advantage afforded by marginal stability could be the limitation of inappropriately persistent and indiscriminate stimulation of this broad spectrum mitogen by promotion of rapid inactivation. Furthermore, partitioning and stabilization by endogenous heparan proteoglycans, most abundant on the basement membranes on which many types of cells divide, could facilitate controlled cellular stimulation consistent with proper tissue architecture. The FGF family of proteins are β-trefoil type structures (Murzin et al., 1992), a structural fold shared by such diverse proteins as the Kunitz type protease inhibitors

[†] This work was supported in part by the Markey Charitable Trust, and by a Florida State University College of Research Creativity grant to M.B.

[‡] The coordinates of the refined structure have been deposited with the Brookhaven National Laboratories Protein Data Bank, accession identification 2AFG.

* To whom correspondence should be addressed.

[§] Florida State University.

[||] Merck Research Laboratories.

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1996.

[1] Abbreviations: aFGF, acidic fibroblast growth factor; bFGF, basic fibroblast growth factor; FGFR, fibroblast growth factor receptor; GuHCl, guanidine hydrochloride; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; F_o, observed structure factor amplitudes; F_c, model structure factor amplitudes; RMS, root mean square; SOS, sucrose octasulfate; hIL-1β, human interleukin-1β; HS, heparin/heparan sulfate; ASU, crystal asymmetric unit.

0006-2960/96/0435-2086$12.00/0

(Blow et al., 1974; Sweet et al., 1974), interleukin-1 (Finzel et al., 1989; Graves et al., 1990), and the muscle protein hisactophilin (Habazettl et al., 1992). The structural fold consists of a three-fold repeat of a four-stranded β-sheet, β-turn motif and is devoid of α-helical structure. A detailed understanding of the biochemical basis of the marginal stability of human aFGF and the surface features available to contribute to heparin and receptor binding requires precise knowledge of its tertiary structure. We report here a refined high-resolution crystal structure of human recombinant aFGF and compare it with the structures of bovine aFGF and human bFGF (Eriksson et al., 1991; Zhang et al., 1991; Zhu et al., 1991).

EXPERIMENTAL PROCEDURES

aFGF Expression, Purification, and Refolding. The 141 amino acid form of human recombinant aFGF was expressed in *Escherichia coli* and purified by sequential chromatography on cation exchange, heparin affinity, and reverse-phase HPLC columns (Linemeyer et al., 1990; Ortega et al., 1991). The resulting protein was greater than 99% pure as determined by SDS−PAGE followed by silver-staining and was stored at −70 °C in the reverse-phase chromatography eluate (7 mM trifluoroacetic acid, 33% acetonitrile). Recombinant aFGF was dialyzed in 6000−8000 M_r cutoff dialysis tubing (Spectra/Por) against two changes of 3 M guanidine hydrochloride (GuHCl, Heico) and 5 mM dithiothreitol (DTT, Calbiochem) for a total of 8 h at 4 °C and refolded by subsequent dialysis versus several changes of crystallization buffer [20 mM NaH_2PO_4, pH 6.20, 10 mM $(NH_4)_2SO_4$, 0.5 mM EDTA, 2 mM DTT] over 30 h at 4 °C. Precipitate was removed by centrifugation at 5000*g* for 5 min. DTT was added to a final concentration of 20 mM, and the solution was filtered through a 0.2 μm syringe filter (Millipore) and concentrated using a Centricon-10 spin concentrator (Amicon). The final aFGF concentration was determined spectrophotometrically using $A_{280} = 1.32$ for a 1 mg/mL solution and diluted in crystallization buffer to 3.0 mg/mL.

Crystallization of aFGF. Pure aFGF solutions were filtered through 0.1 μm spin filters (Millipore) immediately before crystallization. Crystals suitable for X-ray diffraction grew in approximately 4−5 weeks at 4 °C in 10 μL hanging drops by vapor diffusion against 1 mL reservoirs of 9−11% PEG in crystallization buffer.

Space Group Determination and Data Collection. The crystal parameters and space group were determined using precession photography (Charles Supper Co.) with a rotating anode generator (Rigaku RU200). Data collection was accomplished using a Xuong-Hamlin multiwire system. The program STRAT was utilized to optimize the efficiency of data collection (Zhang & Matthews, 1993). Scaling and merging of the data was accomplished using supplied software (San Diego Multi-wire Systems).

Molecular Replacement. The MRCHK software package was used for the self rotation function search and all aspects of molecular replacement rotation and translation searches (Zhang & Matthews, 1994). The search models used included the "A" and "B" structures of the bovine acidic fibroblast growth factor (Zhu et al., 1991). To minimize the inclusion of incorrect information in these search models, those amino acid positions which differ between the bovine and human FGF were first truncated to alanine (a total of 11 residue positions).

The molecular replacement rotation search was optimized by first identifying an appropriate radius for the Patterson vectors (24 Å in this case). After this, the search parameters were optimized for an appropriate resolution shell (10−3.5 Å). Solutions from the rotation function search were applied, and the correctly rotated search models were used in the subsequent translation function search. It was noted that variations in those parameters which led to improvements in the rotation function search in each case led to improvements in the subsequent translation function search. The translation search was also optimized for resolution cutoff (10−4.0 Å), since it was observed that the resolution range which gave the best signal in the rotation function search did not necessarily yield the highest signal in the translation search. Correctly rotated and translated solutions were subsequently included as "known" fragments in further rotation function searches (see Results).

Refinement. Refinement was accomplished using the TNT least-squares refinement package (Tronrud, 1992; Tronrud et al., 1987). Initial rigid body refinement was accomplished using low (10−4 Å) resolution data. Subsequent refinement of individual atomic positions was performed including all low resolution data with incremental addition of high-resolution data. After the inclusion of 2.5 Å data, the bovine/human heterogeneous side chain positions were built into density maps ($2F_o - F_c$ and $F_o - F_c$) using the molecular graphics program O (Jones et al., 1991). After the inclusion of the 2.0 Å data individual atomic positions and thermal factors were simultaneously refined. Refinement of thermal factors included the use of correlated thermal factor restraints (Tronrud et al., 1987). Solvent molecules were included in the final model provided they met the following criteria: (a) there was unambiguous density at a level of 3σ or higher present in density maps, (b) the refined solvent molecules exhibited appropriate hydrogen bonding geometry with a polar atom of the protein (i.e., comprised the first solvent shell), and (c) the refined solvent thermal factors were less than 50 Å^2.

RESULTS

Crystallization, Space Group Determination, and Data Collection. The majority of crystals grew as clusters and were inappropriate for data collection; single crystals were quite rare. In one hanging drop experiment we were able to obtain a single crystal approximately 0.2 × 0.4 × 0.6 mm in size. This single crystal was used for both space group determination and data collection.

The space group was identified as monoclinic ($P2_1$) with real cell parameters $a = 37.1$ Å, $b = 111.7$ Å, $c = 64.0$ Å, and the unique axis $\beta = 90.5°$. It was noted from precession photography of the *h0l* zone that at low resolution systematic absences were present for h = odd indices, indicating a potential 2_1 screw (data not shown). However, this apparent symmetry broke down at higher resolution. Since there was the possibility for a higher order (orthorhombic) symmetry, an attempt was made to merge reflections which would be related by this symmetry. However, the R factor for comparing the structure factor amplitudes for "symmetry" related reflections in this space group was 26%, indicating space group $P2_1$ was correctly assigned. The V_m coefficient

Table 1: Crystal, Data Collection, and Refinement Statistics

Data Collection				
space group		$P2_1$		
cell constants		$a = 37.1$ Å, $b = 111.7$ Å, $c = 64.0$ Å, $\beta = 90.5°$		
V_m (four molecules/ASU)		2.1		
independent reflections		19 981		
highest resolution		2.0 Å		
overall data completeness to 2.0 Å		55%		
R_{merge}		4.1%		
Wilson temperature factor		22.5 Å^2		
shell lower limit (Å)	$I/\sigma I$	reflections possible	reflections collected	% shell completeness
3.22	28.9	8528	6102	72%
2.55	8.5	8449	5227	62%
2.23	3.4	8431	4023	48%
2.03	2.1	8384	3553	42%
Refinement				
crystallographic residual			16.9%	
RMS bond length deviation			0.009 Å	
RMS bond angle deviation			1.9°	

Table 2: Self-Rotation Function Peaks[a]

angle (deg)			
ϕ	φ	κ	peak
89.6	90.0	180.0	95.6
0.0	0.0	180.0	70.5
0.0	90.3	180.0	68.7

[a] Values are given in polar angles (ϕ, φ, κ), and peak heights are set relative to the identity peak value of 100. The resolution range is from 10 to 4 Å. The three peaks of the self-rotation function are nearly parallel to the *b*, *c*, and *a* cell edges, respectively. The first peak represents the crystallographic 2_1 screw along the *b* cell edge. The second peak represents a local (noncrystallographic) screw parallel to the *c* cell edge which relates molecules A and D as a pair with molecules B and C in the asymmetric unit. The third peak represents a pseudo 2_1 screw along the *a* cell edge, apparent at low resolution but breaking down at higher resolution, which relates molecule A with B as well as C with D.

(Matthews, 1968) was 2.1 with four molecules estimated in the asymmetric unit. During space group determination the crystal was positioned such that approximately half of the crystal was irradiated. After this the crystal and goniometer were transferred to a Xuong-Hamlin multiwire system, and the assembly was translated so that the region which had not been irradiated during the space group determination was used for data collection. Data between 20 and 2 Å resolution were collected for approximately 30 h. At this time the signal to noise ratio for the intensities of all reflections within the 2.2−2.0 Å resolution shell fell below 2.0σ, and there was a drastic crystal decay. Data collection was halted with an overall completeness of 55% from 20 to 2.0 Å and 42% complete for the 2.2−2.0 Å resolution shell. Crystal and data collection statistics are summarized in Table 1.

Molecular Replacement. Results of a self-rotation function search indicated the presence of a set of essentially mutually orthogonal two-fold axes parallel to the major crystallographic axes (Table 2). In the initial molecular replacement rotation function search, with the "A" model of bovine aFGF (using data from 10 to 3.5 Å), a peak approximately 9σ above background was identified. This was consistent with the top peak found using the "B" model of bovine aFGF as a search model. Lesser peaks were inconsistent between the two search models. However, in each case a lesser peak could be found which was related to the top peak by a 180° rotation parallel to either the *a* or *c* cell edge. Thus, from the self-rotation function results, this top peak was tentatively identified as a correct rotation function solution and additional potential solutions were identified from among the lesser peaks.

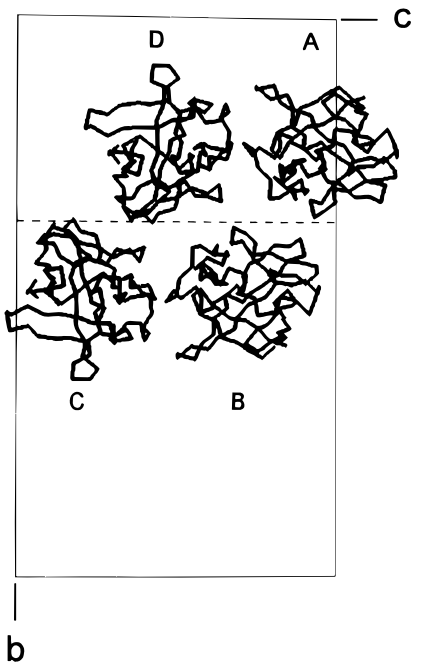

FIGURE 1: View down the *a* cell edge showing the arrangement of the four aFGF molecules (A, B, C, and D) in the asymmetric unit. The contents of the asymmetric unit can be described as a pair of "dimers". Molecules A and D as a pair are related to molecules B and C by a local screw axis parallel to the *c* cell edge (shown with a dashed line). The screw angle is approximately 180° and the screw distance is 15.9 Å, or approximately one quarter of the *c* cell edge.

The top peak of the bovine aFGF A search model (after application of the rotation function solution) yielded a 7σ peak on the translation function search, suggesting the model had been correctly placed in the crystal cell. This new model was therefore included as a "known" fragment in a subsequent rotation function search. In this case, the subsequent top rotation function solution was consistent between the bovine aFGF A and B search models. By this iterative process, four molecules were placed in the unit cell.

The four rotationally and translationally positioned molecules (all derived from the bovine aFGF A model) packed together with no significant overlapping intermolecular contacts. The *R* factor calculation of this initial structure was 37%, suggesting that the overall placement of these molecules in the unit cell was correct. The four molecules were subsequently given the chain identifications A, B, C, and D. The relationship between the four molecules in the asymmetric unit can be described as a pair of "dimers". Molecules A and D as a pair are related to molecules B and C by a local nonintegral screw axis parallel to the *c* cell edge (Figure 1). The rotation angle is approximately 180° with a screw distance of 15.9 Å, or very nearly one quarter of the *c* cell edge distance. From the arrangement of the molecules in the unit cell, and crystallographic symmetry operators, molecule A is related to B by a 2_1 screw relationship parallel to the *a* cell edge (likewise with molecules D and C). Since the molecules are independent and not identical, this relationship is apparent at low

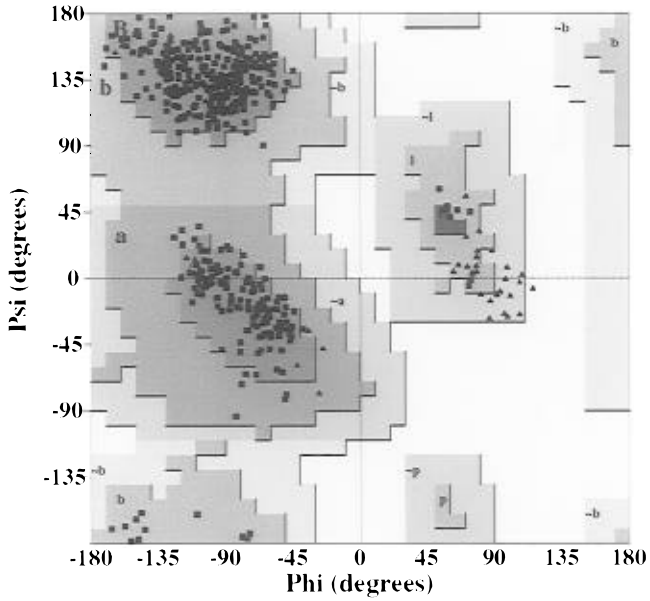

FIGURE 2: Plot of the main chain ϕ,ψ angles (Ramachandran plot) for the four independent structures of human aFGF. A total of 513 residue positions are plotted, with 48 glycine residues shown as triangles. Three hundred ninety-six residues (91%) are within most favored regions (*A*, *B*, and *L*), and 39 residues (9%) are within additional allowed regions (*a*, *b*, *l*, and *p*) (Laskowski et al., 1993).

resolution but breaks down at higher resolution (around 3.7 Å). Therefore, with regard to the peaks observed in the self-rotation function (Table 2) the first peak represents the crystallographic 2_1 screw along the *b* cell edge. The second peak represents a local (noncrystallographic) screw parallel to the *c* cell edge which relates molecules A and D as a pair with molecules B and C in the asymmetric unit. The third peak represents a pseudo 2_1 screw along the *a* cell edge, apparent at low resolution but breaking down at higher resolution, which relates molecule A to B and likewise D to C.

Refinement. Rigid body refinement resulted in a 1% reduction in the *R* factor of the model. After inclusion of the 2.5 Å data, various solvent groups were built into the structure (see below). After inclusion of all data to 2.0 Å resolution correlated *B* factor restraints were included simultaneously with *xyz* positional refinement. The *R* factor of the model at this point was approximately 17.5%. To confirm their position, side chain residues were deleted in groups representing approximately 5% of the total number of atoms in the structure. The molecule was subsequently refined, and "omit" density maps ($F_o - F_c$) were calculated. The deleted side chains were rebuilt into the new difference density maps. At the end of this procedure all of the solvent inaccessible residues and most of the surface accessible residues had been reevaluated. Subsequently, the overall main chain geometry statistics remained relatively unchanged; however, the side chain rotamer distributions were more consistent with highly refined structures (Laskowski et al., 1993), and the *R* factor reduced to 16.9%. The final refinement statistics are listed in Table 1. The Ramachandran plot for the refined structure is shown in Figure 2.

No electron density was observed for the first nine or ten residues, or the last one or two residues in each molecule of aFGF. Residues 10−138 were visible in the A molecule, residues 10−138 in B, residues 11−137 in C, and residues 11−138 in D. The electron density was contiguous over the length of the polypeptide chain(s). Side chain orientations for virtually all of the residue positions were determined with exceptions being residues Lys A10, Gln A40, Ser A138, Lys B10, Ser B47, Ser B138, Ser C47, Lys D40, Ser D47, Ser D50, Val D137, and Ser D138. Overall, the main chain superpositions for the four independent structures are quite similar (Figure 3). The greatest deviations are for the loop regions around residues 18, 50, and 92. The root mean square (RMS) deviation for main chain atoms in the different structures varies between 0.29 and 0.52 Å (Table 3). The A and D molecules as a pair overlay with the B and C molecules with an RMS deviation of 0.50 Å for main chain atoms. In a comparison with the structures of human bFGF and bovine aFGF, the A structure of bovine aFGF is most similar to the four structures of human aFGF (Table 3).

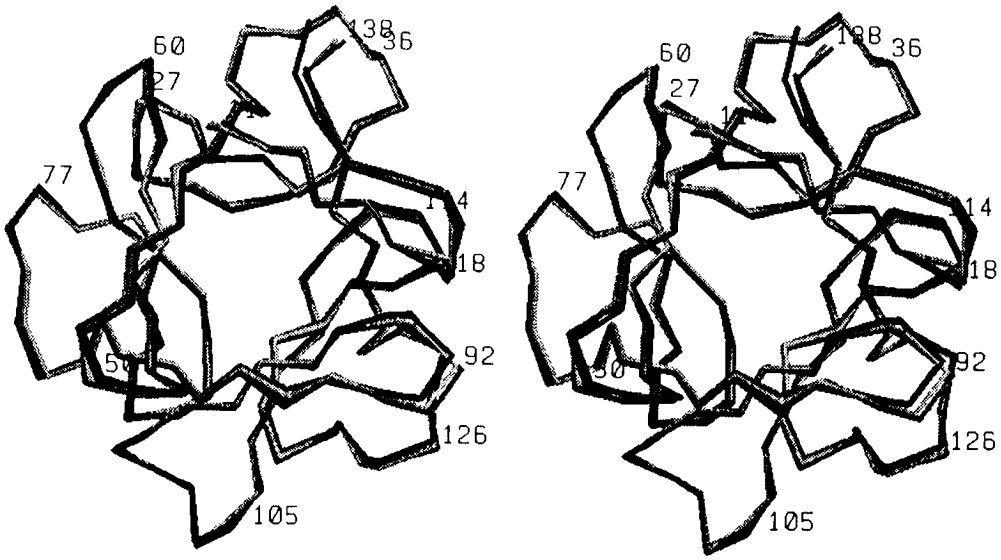

FIGURE 3: Superposition of the main chain (Cα) trace of the four molecules of human aFGF overlaid with one another. Molecule A is represented as light gray, and molecules B−D are a progressively darker shade. The main chain atom RMS deviation between the different molecules varies between 0.29 and 0.52 Å (Table 3). The areas of greatest deviation in atomic position are located at the surface loops near residues 18, 50, and 92.

Table 3: Comparison of FGF Structures[a]

	RMS Deviation (Å)			
molecule	A	B	C	D
A		0.29	0.42	0.51
B			0.45	0.52
C				0.43
D				
human bFGF	0.66	0.65	0.73	0.71
bovine aFGF (A)	0.51	0.52	0.55	0.55
bovine aFGF (B)	0.67	0.64	0.72	0.74

[a] Root mean square (RMS) deviation of main chain atoms (positions 11−137) for a comparison of the four independent molecules (A−D) of human aFGF in the crystal asymmetric unit. The human aFGF structures are also compared to the human bFGF structure (Eriksson et al., 1991) and the bovine aFGF structure(s) (Zhu et al., 1991).

Table 4: Sulfate Ion Hydrogen Bonding Interactions (Å)[a]

	Asn 18 Nδ_2	Lys 113 N	Lys 118 Nζ	Sol 208
SO_4 A 150	3.1	3.1		2.3
SO_4 B 150	3.1	2.7	3.1	2.9
SO_4 C 150	2.8		2.8	
SO_4 D 150	3.2	2.8	2.9	

[a] Hydrogen bonding interactions are listed if they are less than 3.4 Å in distance and have a acceptor−donor angle of >90° and <180°.

Addition of Solvent Groups. Prior to the addition of solvent groups, the largest area of difference density ($F_o - F_c$) corresponded to regions on the surface of each of the four molecules adjacent to residues Asn 18, Lys 113, and Lys 118 (Figure 4). These residues have been implicated in heparin binding (Thompson et al., 1994), and in the binding of sulfate ions in the related human bFGF structure (Eriksson et al., 1991, 1993). Since the crystallization buffer contained $(NH_4)_2SO_4$ (see above), this positive density was interpreted as bound sulfate ion (Figure 4). In all four molecules the interaction with sulfate ion primarily involves the atoms of Asn 18 Nδ_2, Lys 113 N, and Lys 118 Nζ as well as a local solvent molecule (Table 4). Sulfate ions associated with molecules A and B are in general more ordered than in molecules C and D.

In addition to the four sulfate ions, a total of 52 water molecules were added to the structure. Fifteen of these are associated with molecule A, 19 with B, 9 with C, and 9 with D. Solvent molecules have been numbered to reflect related positions. Five solvent molecules (200−204) are conserved at equivalent positions in all four structures. One solvent molecule (205) is found in three structures (A, B, and C), and several solvent molecules (206−210) can be found at equivalent positions in more than one structure. A number of positive difference density features were present within the core regions of the aFGF structures and were interpreted as internally bound solvent groups. Of the 52 added water molecules, 29 are completely buried in the protein structures (9 in A, 8 in B, 7 in C, and 5 in D). The highly conserved solvent molecules 200−205 are all at solvent inaccessible positions.

DISCUSSION

Sulfate Ions. The bound sulfate ions interact in each case with the Nδ_2 atom of side chain Asn 18 (Table 4). Hydrogen bonding interactions are observed for three of the four sulfate ions (A, B, and D) with the main chain amide group of residue Lys 113. Hydrogen bonding interactions are also observed for three of the four sulfate ions (B, C, and D) with the Nζ atom of Lys 118. These data suggest that the interaction with the Nδ_2 atom of the side chain of Asn 18 is the most conserved interaction. A similar bound sulfate has been observed in human bFGF (Eriksson et al., 1991). Equivalent atomic positions involved in the binding of sulfate ions are observed in both structures. The sulfate ions associated with molecules A and B appear more ordered (as judged by greater electron density and lower refined thermal factors) than the ions associated with the C or D molecules. The A and B sulfate ions have an additional hydrogen bonding partner in solvent 208 (Table 4). In turn, this solvent molecule hydrogen bonds (in both cases) with a different molecule in the unit cell. Thus, the more highly ordered A and B sulfate ions may be due in part to this intermolecular contact.

The sulfate ion binding site has been postulated to contribute to the heparin binding site in both aFGF and bFGF (Eriksson et al., 1991). A crystal structure of bovine aFGF complexed with the heparin analog sucrose octasulfate (SOS) has been reported (Zhu et al., 1993). The primary contacts of SOS involve residues Lys 112, Arg 116, Lys 118, and

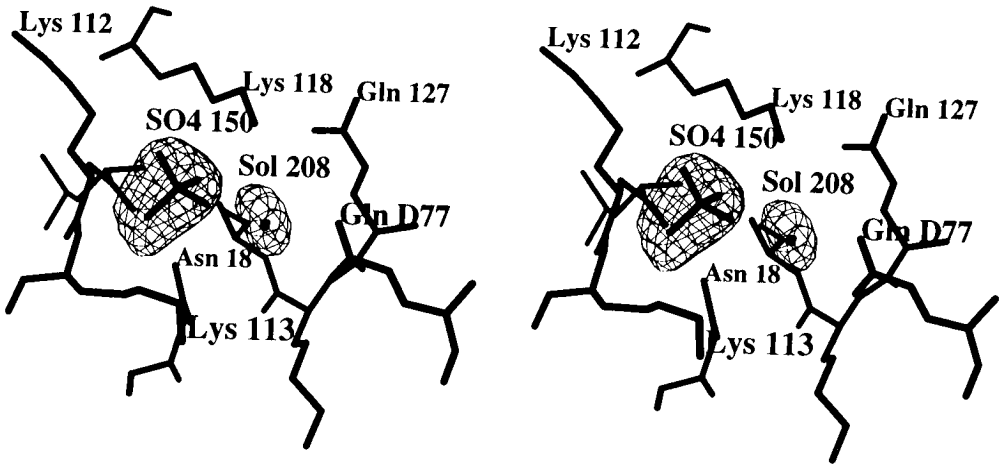

FIGURE 4: Structure of human aFGF (molecule A) in the region of residues Asn 18, Lys 113, and Lys 118. Also shown is an $F_o - F_c$ difference density map contoured at $+4\sigma$. A sulfate ion (residue 150) has been modeled into the large area of positive difference density adjacent to Lys 112. The sulfate ion interacts primarily with the Nδ_2 atom of Asn 18, the main chain amide of Lys 113, and the Nζ atom of Lys 118 (Table 4). The smaller area of positive difference density adjacent to Lys 118 has been modeled as a water molecule (Sol 208).

Arg 122. Asn 18 does not appear to play a significant role in the binding of SOS to bovine aFGF. In alanine scanning mutagenesis studies of residues surrounding the sulfate binding site of human bFGF, Springer and co-workers (Thompson et al., 1994) identified positions Asn 18, Lys 113, Lys 118, and Gln 127 (aFGF numbering scheme) as contributing to heparin binding. Gln 127 is not observed to interact with any of the four bound sulfate ions described here. Additionally, the interaction of sulfate ion with position 113 involves the main chain amide rather than the side chain Nζ atom. Therefore, the interaction of side chains Lys 113 and Gln 127 in the binding of heparin may involve interactions separate from those directly interacting with the sulfate moiety.

Ordered Solvent. There are five ordered solvent molecules (residues 200−204) observed at equivalent positions in each of the four independent structures of human aFGF. Each of these are at solvent-inaccessible locations in the structure. Three of these (solvent residues 200−202) are related by an internal pseudo-three-fold axis of symmetry which is present in β-trefoil type structures such as FGF (McLachlan, 1979; Murzin et al., 1992). In each case, these solvent groups hydrogen bond with two main chain carbonyl groups and a single main chain amide group. Thus, these solvent residues form a hydrogen bonding network that connects three different secondary structure elements. Solvent 203 forms hydrogen bonds with the main chain amide of Gly 110 and main chain carbonyl groups of Phe 108 and Lys 118. Solvent 204 forms hydrogen bonds with the main chain amide of Gly 19, the main chain carbonyl of Lys 112, and the Nδ_1 atom of His 21. Thus, in each case, three hydrogen bonding partners for the solvent groups are observed, with the most common arrangement involving two main chain carbonyl groups (as hydrogen bond acceptors) and a main chain amide group (as a hydrogen bond donor). The hydrogen bonding partners of these solvent groups are themselves solvent inaccessible. Furthermore, the only hydrogen bonding partners of these protein groups are the solvent molecules. Since buried polar groups with unsatisfied hydrogen bonding arrangements contribute to the destabilization of protein structures (Blaber et al., 1993; Harpaz et al., 1994), these buried water molecules appear to be an integral part of the FGF structure and likely contribute to overall stability.

Cysteine Residues. There are three cysteine residues in aFGF at positions 16, 83, and 117. Cysteine residues 16 and 83 are conserved between acidic and bFGF, whereas cysteine 117 is unique to aFGF. While many extracellular proteins contain cysteine residues which form disulfide bonds, thereby contributing to overall stability, aFGF has been shown previously to contain only free cysteine residues (Linemeyer et al., 1990).

In the absence of reducing agents, aFGF has been demonstrated to undergo rapid, proteolytic independent inactivation (Ortega et al., 1991). This inactivation has been attributed to the formation of intra- or interchain disulfide bonds, or the formation of mixed thiols. All three cysteine residues in the four independent structures of aFGF are located at solvent-inaccessible positions. Thus, oxidation of these cysteine residues would require structural changes allowing accessibility. Conversely, in the absence of side chain or other conformational changes, formation of mixed thiols would be incompatible with the native structure—the adducts would have to be accommodated within the core region of the protein.

Table 5: Central Cavity Volume (Å^3) of Human aFGF[a]

	probe radius (Å)		
molecule	1.4	1.3	1.2
A	21.4	42.3	48.3
B	0	0	20.3
C	23.6	34.9	40.1
D	0	22.6	27.6

[a] The central cavity volume for the four independent molecules (A−D) of human aFGF. Cavity volumes are calculated (Connolly, 1993) using the indicated probe radii and with all bound solvent molecules included.

The structure of bFGF has been solved with cysteines at positions 87 and 92 (bFGF numbering scheme) each covalently linked to β-mercaptoethanol (Eriksson et al., 1991). Of these, cysteine 92 is conserved in aFGF (position 83). Cysteine 83 in aFGF has a *gauche*$^+$ side chain χ_1 torsion angle. In the bFGF/β-mercaptoethanol adduct structure the cysteine side chain has rotated to give a χ_1 torsion angle of *gauche*$^-$, thus positioning the β-mercaptoethanol group toward the solvent-exposed exterior. Cysteine 117 in aFGF adopts a *trans* χ_1 torsion angle (in all four molecules); solvent accessibility for a potential disulfide adduct would be possible if it adopted the *gauche*$^-$ conformation. Unlike Cys 83 or 117, cysteine 16 cannot achieve solvent accessibility for any side chain rotamer. Therefore, of all the free cysteine residues, a disulfide adduct at position 16 would be postulated to produce the most destabilizing effects upon aFGF.

A Cys to Ser mutant at position 16 of aFGF, as well as the equivalent substitution in bFGF, has been characterized (Ortega et al., 1991; Seno et al., 1988). In both cases the specific activity of the mutant protein was approximately half that of the wild-type protein. Therefore, while the substitution to Ser may have diminished mixed thiol mediated irreversible denaturation, the introduction of a Ser residue into the core of the protein appears to have significantly destabilized the protein.

Internal Cavities. Of particular interest is the presence and nature of internal cavities within the structure of aFGF. The related β-trefoil type protein human interleukin-1β (hIL-1β) has a centrally located cavity with a volume of approximately 100 Å^3 which has been reported to contain positionally disordered solvent (Ernst et al., 1995). Cavity calculations (Connolly, 1993) for the four independent molecules of human aFGF, using various probe radii, are listed in Table 5. Molecule B contains no apparent central cavity until a probe radius of 1.2 Å is reached, at which point a small central cavity of approximately 20 Å^3 is identified. Molecule A exhibits the largest internal cavity, which approaches some 48 Å^3 with a 1.2 Å radius probe. Molecules C and D are intermediate between these values.

The residues which define the boundary of the central cavity include Leu 14, Leu 23, Leu 44, Ile 56, Leu 65, Phe 85, Tyr 97, Val 109, and Phe 132 (Figure 5). There are no major differences in rotamer orientations for this set of residues between the different aFGF structures, and the variation in cavity sizes is the result of a combination of main chain and side chain positional shifts. The RMS deviation for the atoms of this set of residues between the

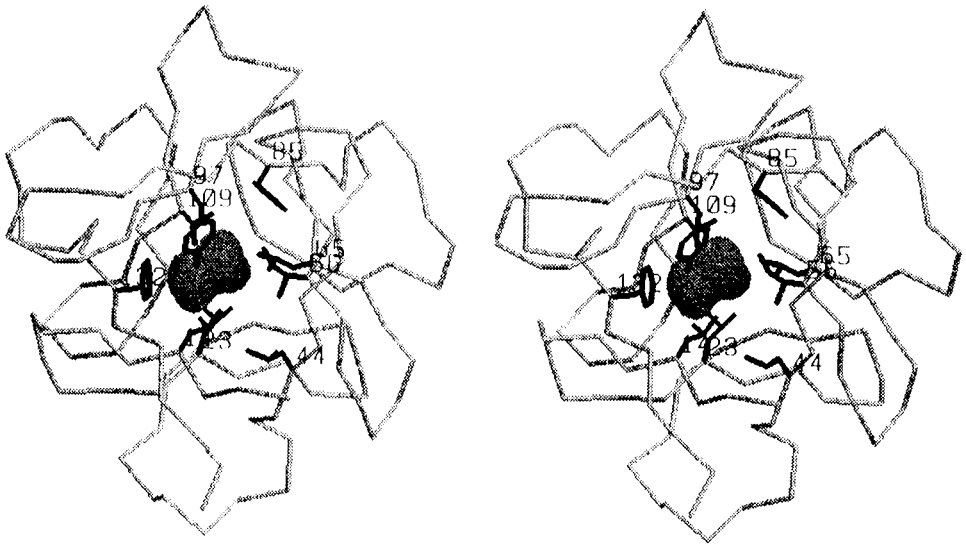

FIGURE 5: Main chain (Cα) trace of human aFGF molecule A. The central cavity [approximately 48 Å^3; calculated using a 1.2 Å probe radius (Connolly, 1993)] is shown as a "dot" surface. Also shown are the residue positions (Leu 14, Leu 23, Leu 44, Ile 56, Leu 65, Phe 85, Tyr 97, Val 109, and Phe 132) which define the central cavity.

four aFGF molecules varies between 0.30 and 0.44 Å, similar to that observed for the entire set of main chain atoms (Table 3). Residues Leu 14 and Leu 23 exhibit the greatest relative movement between the four structures, ranging up to approximately 1.5 Å. The cavity calculations are clearly sensitive to the choice of probe size (Table 5) and therefore, also to slight positional shifts of the centrally located residues. Thus, the variation in cavity sizes for the four molecules of aFGF may reflect a combination of crystal packing interactions and small atomic positional error.

The evaluation of the four independent structures of aFGF indicates that compared with IL-1β the central cavity is significantly smaller. No crystallographically ordered solvent was visible in any of the central cavities. The set of residues defining the central cavity is conserved between human aFGF, bovine aFGF, and human bFGF but not with IL-1β. The internal cavity volumes for human bFGF and bovine aFGF, between 15 and 21 Å^3 (1.2 Å probe radius), are similar to that of human aFGF. An internal cavity in human aFGF between 20 and 48 Å^3 would be expected to contribute to the destabilization of the overall structure by some 0.6−1.5 kcal/mol (Eriksson et al., 1992).

Structural Determinants of Receptor Interactions. FGF receptors are transmembrane tyrosine kinases containing either two or three extracellular immunoglobulin-like domains that contain the FGF binding sites. Four FGFR genes have been identified, three of which (FGFR1−3) can undergo alternative splicing of the third exon which encodes the C-terminal half of the membrane-proximal Ig-like domain. Although aFGF and bFGF are 55% identical in amino acid sequence and have very similar overall tertiary structures, they do not bind equivalently to all known FGF receptors. Inclusion of exon IIIc generates transmembrane receptors that bind aFGF and bFGF with comparable affinity. In contrast, alternative incorporation of exon IIIb encodes receptors that retain high affinity binding for aFGF but form weaker complexes with bFGF (Chellaiah et al., 1994).

Interactions of bFGF and FGF receptor 1 incorporating exon IIIc (FGFR1/IIIc) have been demonstrated by ultracentrifugation to involve higher and lower affinity bFGF receptor binding sites (Pantoliano et al., 1994). This interaction, which results in the formation of a bFGF/FGFR1 complex with 1:2 stoichiometry, includes heparin/heparan sulfate (HS) forming a ternary complex. Mutagenesis studies of bFGF have also identified two distinct binding domains for FGFR1 (Springer et al., 1994). A primary higher affinity site is comprised of residues Tyr 15, Arg 35, Asn 92, Tyr 94, Leu 133, and Met 135 (aFGF numbering scheme). A second lower affinity site (K_d ~250-fold higher) on the opposite side of the molecule is comprised of residues Lys 101, Tyr 102, and Trp 107. The identification of two binding sites on FGF has suggested a mechanism whereby receptor dimerization and subsequent signal transduction can be mediated by a single FGF molecule (Springer et al., 1994).

The residues which define the lower affinity receptor binding site of bFGF include a five-residue surface loop between the eighth and ninth β strands of the structure, plus Trp 107—the N-terminal residue of the ninth β strand. The corresponding region in aFGF is comprised of a seven-residue loop from positions Lys 101 to His 106 and the adjacent Trp 107, which is conserved between the two FGFs (Figure 6a). This longer loop region from aFGF has been substituted into bFGF, and its effects on receptor binding have been characterized (Seddon et al., 1995). Basic FGF, and the mutant containing the low affinity receptor binding region from aFGF, bound with similar affinities to the exon IIIc forms of FGFR1 and FGFR2. However, only aFGF and the bFGF mutant containing the low affinity receptor binding region from aFGF bound to the keratinocyte growth factor receptor, the exon IIIb-containing version of FGFR2 (Seddon et al., 1995). These results indicate that the lower affinity receptor binding site is a critical determinant for not only receptor dimerization but also selective receptor recognition of the exon IIIb-containing FGFR2.

The high-resolution structure of human aFGF presented here allows a detailed comparison with the receptor binding regions of human bFGF. With the exception of a Leu substitution for Met at position 135, the residues comprising the high affinity FGFR binding site are conserved in aFGF. These residues in aFGF essentially superimpose on the corresponding residues of bFGF (Eriksson et al., 1991), indicating that receptor interactions of this FGF surface

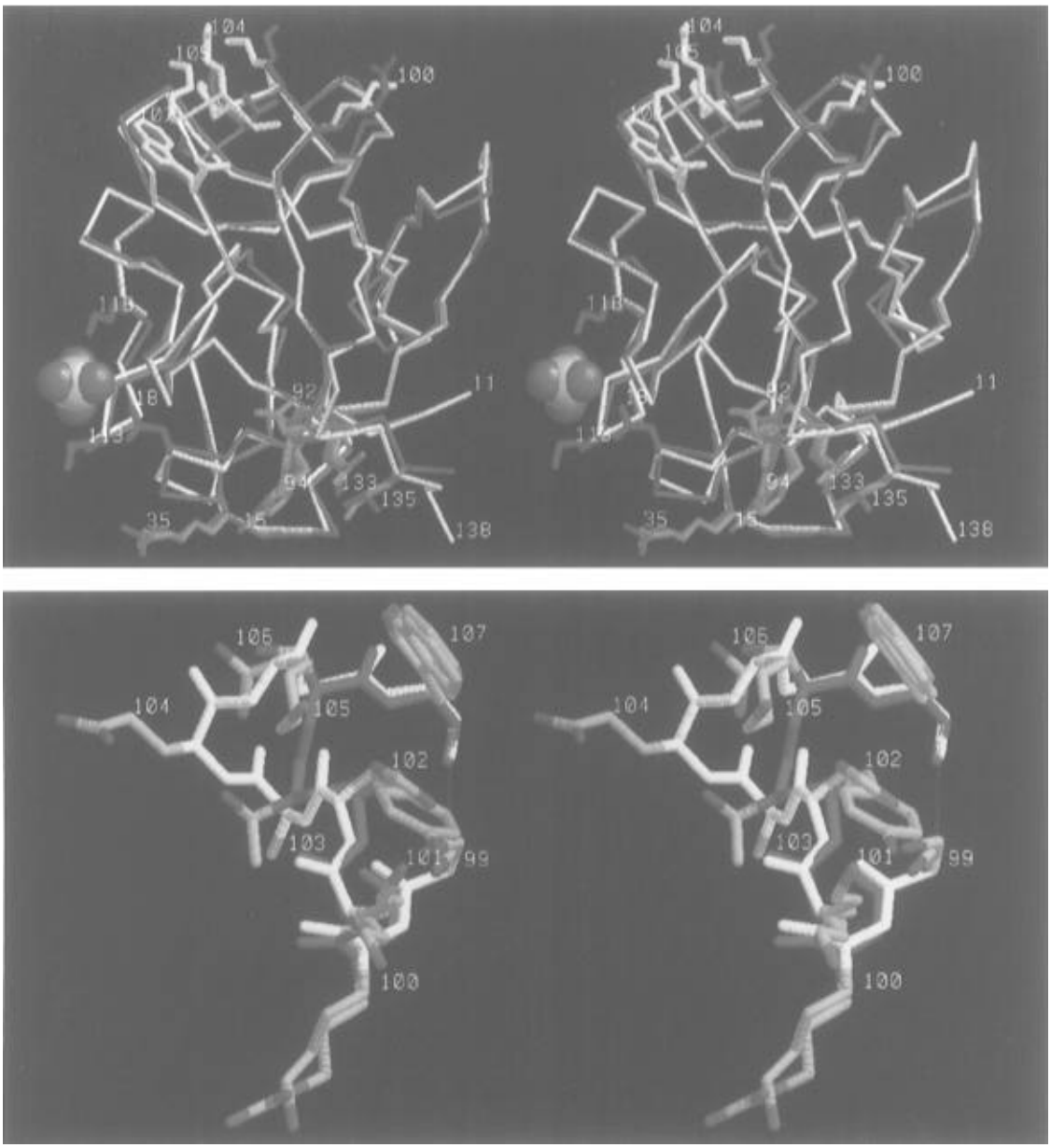

FIGURE 6: (a, top) Main chain (Cα) trace of human aFGF molecule A (white) overlaid with the main chain trace of human bFGF (Eriksson et al., 1991). Shown are the residue positions (Tyr 15, Arg 35, Asn 92, Tyr 94, Leu 133, and Leu 135) associated with the high affinity FGF receptor site (aFGF in green, bFGF in cyan). A similar comparison of residue positions defining the low affinity FGF receptor binding site (residues Lys 101 to Trp 107 in aFGF) is shown (aFGF in yellow, bFGF in red). The loop defining the low affinity binding site in bFGF is shorted by two residue positions in comparison with aFGF. The location of the bound sulfate ion in human aFGF is shown (spacefilling representation) as are the residues defining the sulfate binding site (purple). (b, bottom) Close up view of the loop region 99−107 of human aFGF (backbone in yellow) overlaid with the corresponding region of human bFGF (backbone in red). The main chain hydrogen bond from N 99 to O 107 is the initial hydrogen bond defining local β-sheet secondary structure. The amino acid sequence of human aFGF from position 99 to 107 includes Ser, Lys, Lys, His, Ala, Glu, Lys, Asn, and Trp. The corresponding loop of human bFGF is two residues shorter and is comprised of residues Ser, Arg, Lys, Tyr, Thr, Ser, and Trp.

would be expected to be very similar for these two FGF ligands and would offer little discrimination for receptor subtype binding (Figure 6a).

The lower affinity receptor binding site ($\beta8-\beta9$ loop) differs in both length and sequence between aFGF and bFGF. The loop conformations of the five-residue sequence in bFGF (Arg-Lys-Tyr-Thr-Ser) and the corresponding seven-residue sequence (residues 100−106; Lys-Lys-His-Ala-Glu-Lys-His) in aFGF are significantly different (Figure 6b) providing unique structural determinants for receptor recognition. Thus, the structural data support the hypothesis that the high affinity site which provides the bulk of the receptor binding energy is structurally conserved between bFGF and aFGF, whereas the structurally divergent low affinity site functionally determines receptor subtype specificity.

The lower affinity receptor binding loop that confers discriminatory receptor recognition between aFGF and bFGF is also the loop of greatest variability in amino acid length among the nine known FGFs. While FGF-4 (HST/KFGF), FGF-6, and FGF-8 have the same length loop as bFGF, FGF-2 (Int-2), FGF-5, FGF-7 (KGF), and FGF-9 have insertions of 16, 6, 4, and 4 amino acid residues, respectively (Miyamoto et al., 1993; Tanaka et al., 1992; Thomas, 1993). Thus, variability in this loop is likely not only to contribute to receptor binding discrimination between aFGF and bFGF but also among other members of the FGF family.

ACKNOWLEDGMENT

We thank Drs. Larry Weaver, Dale Tronrud, Tim Logan, and Michael Chapman for helpful discussions and Dr. Brian Matthews for his support and encouragement in the development of this project.

REFERENCES

Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W., & Matthews, B. W. (1993) *Biochemistry 32*, 11363−11373.
Blow, D. M., Janin, J., & Sweet, R. M. (1974) *Nature 249*, 54−57.
Burgess, W. H., & Maciag, T. (1989) *Annu. Rev. Biochem. 58*, 575−606.
Chellaiah, A. T., McEwen, D. G., Werner, S., Xu, J., & Ornitz, D. M. (1994) *J. Biol. Chem. 269*, 11620−11627.
Connolly, M. L. (1993) *J. Mol. Graph. 11*, 139−141.
Copeland, R. A., Ji, H., Halfpenny, A. J., Williams, R. W., Thompson, K. C., Herber, W. K., Thomas, K. A., Bruner, M. W., Ryan, J. A., Marquis-Omer, D., Sanyal, G., Sitrin, R. D., Yamazaki, S., & Middaugh, C. R. (1991) *Arch. Biochem. Biophys. 289*, 53−61.
Eriksson, A. E., Cousens, L. S., Weaver, L. H., & Matthews, B. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3441−3445.
Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178−183.
Eriksson, A. E., Cousens, L. S., & Matthews, B. W. (1993) *Protein Sci. 2*, 1274−1284.
Ernst, J. A., Clubb, R. T., Zhou, H.-X., Gronenborn, A. M., & Clore, G. M. (1995) *Science 267*, 1813−1817.
Finzel, B. C., Clancy, L. L., Holland, D. R., Muchmore, S. W., Watenpaugh, K. D., & Einspahr, H. M. (1989) *J. Mol. Biol. 209*, 779−791.
Gospodarowicz, D., & Cheng, J. (1986) *J. Cell. Physiol. 128*, 475−484.
Graves, B. J., Hatada, M. H., Hendrickson, W. A., Miller, J. K., Madison, V. S., & Satow, Y. (1990) *Biochemistry 29*, 2679−2684.
Habazettl, J., Gondol, D., Wiltscheck, R., Otlewski, J., Schleicher, M., & Holak, T. A. (1992) *Nature 359*, 855−858.
Harpaz, Y., Elmasry, N., Fersht, A. R., & Henrick, K. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 311−315.
Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Linemeyer, D. L., Kelly, L. J., Menke, J. G., Gimenez-Gallego, G., DiSalvo, J., & Thomas, K. A. (1987) *Biotechnology 5*, 960−965.
Linemeyer, D. L., Menke, J. G., Kelly, L. J., DiSalvo, J., Soderman, D., Schaeffer, M.-T., Ortega, S., Gimenez-Gallego, G., & Thomas, K. A. (1990) *Growth Factors 3*, 287−298.
Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
McLachlan, A. D. (1979) *J. Mol. Biol. 133*, 557−563.
Miyamoto, M., Naruo, K.-I., Seko, C., Matsumoto, S., Kondo, T., & Kurokawa, T. (1993) *Mol. Cell. Biol. 13*, 4251−4259.
Murzin, A. G., Lesk, A. M, & Chothia, C. (1992) *J. Mol. Biol. 223*, 531−543.
Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E, & Leder, P. (1992) *Mol. Cell. Biol. 12*, 240−247.
Ortega, S., Schaeffer, M.-T., Soderman, D., DiSalvo, J., Linemeyer, D. L., Gimenez-Gallego, G., & Thomas, K. A. (1991) *J. Biol. Chem. 266*, 5842−5846.
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.
Rapraeger, A. C., Krufka, A., & Olwin, B. B. (1991) *Science 252*, 1705−1708.
Roghani, M., Mansukhani, A., Dell'Era, P., Bellosta, P., Basilico, C., Rifkin, D. B., & Moscatelli, D. (1994) *J. Biol. Chem. 269*, 3976−3984.
Rosengart, T. K., Johnson, W. V., Friesel, R., Clark, R., & Maciag, T. (1988) *Biochem. Biophys. Res. Commun. 152*, 432−440.
Seddon, A. P., Aviezer, D., Li, L.-Y., Bohlen, P., & Yayon, A. (1995) *Biochemistry 34*, 731−736.
Seno, M., Sasada, R., Iwane, M., Sudo, K., Kurokawa, T., Ito, K., & Igarashi, K. (1988) *Biochem. Biophys. Res. Commun. 151*, 701−708.
Springer, B. A., Pantoliano, M. W., Barbera, F. A., Gunyuzlu, P. L., Thompson, L. D., Herblin, W. F., Rosenfeld, S. A., & Book, G. W. (1994) *J. Biol. Chem. 269*, 26879−26884.
Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., & Blow, D. M. (1974) *Biochemistry 13*, 4212−4228.
Tanaka, A., Miyamoto, K., Minamino, N., Takeda, M., Sato, B., Matsuo, H., & Matsumoto, K. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8928−8932.
Thomas, K. A. (1993) *Neurotrophic Factors* (Loughlin, S. E., & Fallon, J. H., Eds.) pp 285−312, Academic Press, Inc., San Diego.
Thompson, L. D., Pantoliano, M. W., & Springer, B. A. (1994) *Biochemistry 33*, 3831−3840.
Tronrud, D. E. (1992) *Acta Crystallogr. A48*, 912−916.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., & Ornitz, D. M. (1991) *Cell 64*, 841−848.
Zhang, J., Cousens, L. S., Barr, P. J., & Sprang, S. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3446−3450.
Zhang, X.-J., & Matthews, B. W. (1993) *J. Appl. Crystallogr. 26*, 457−462.
Zhang, X.-J., & Matthews, B. W. (1994) *Acta Crystallogr. D50*, 675−686.
Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G. M., Arakawa, T., Hsu, B. T, & Rees, D. C. (1991) *Science 251*, 90−93.
Zhu, X., Hsu, B. T., & Rees, D. C. (1993) *Structure 1*, 27−34.

BI9521755