Correlation between the 1.6 Å crystal structure and mutational analysis of Keratinocyte Growth Factor

TIMOTHY D. OSSLUND,¹ RASHID SYED,¹ ELIZABETH SINGER,¹ ERIC W.-J. HSU,¹
REBECCA NYBO,¹ BAO-LU CHEN,² TIMOTHY HARVEY,¹ TSUTOMU ARAKAWA,¹
LINDA OWERS NARHI,¹ ARTHUR CHIRINO,¹ AND CHARLES F. MORRIS¹

¹Amgen, Amgen Center, Thousand Oaks, California, 91320-1789

²Chiron Corporation, 4560 Horton Street, Emeryville, California 94608

(RECEIVED November 24, 1997; ACCEPTED May 15, 1998)

Abstract

A comprehensive deletion, mutational, and structural analysis of the native recombinant keratinocyte growth factor (KGF) polypeptide has resulted in the identification of the amino acids responsible for its biological activity. One of these KGF mutants (Δ 23KGF-R144Q) has biological activity comparable to the native protein, and its crystal structure was determined by the multiple isomorphous replacement plus anomalous scattering method (MIRAS). The structure of KGF reveals that it folds into a β -trefoil motif similar to other members of fibroblast growth factor (FGF) family whose structures have been resolved. This fold consists of 12 anti-parallel β -strands in which three pairs of the strands form a six-stranded beta-barrel structure and the other three pairs of β -strands cap the barrel with hairpin triplets forming a triangular array. KGF has 10 well-defined beta strands, which form five double-stranded anti-parallel beta-sheets. A sixth poorly defined β -strand pair is in the loop between residues 133 and 144, and is defined by only a single hydrogen bond between the two strands. The KGF mutant has 10 additional ordered amino terminus residues (24–33) compared to the other FGF structures, which are important for biological activity. Based on mutagenesis, thermal stability, and structural data we postulate that residues TRP125, THR126, and His127 predominantly confer receptor binding specificity to KGF. Additionally, residues GLN152, GLN138, and THR42 are implicated in heparin binding. The increased thermal stability of Δ 23KGF-R144Q can structurally be explained by the additional formation of hydrogen bonds between the GLN side chain and a main-chain carbonyl on an adjoining loop. The correlation of the structure and biochemistry of KGF provides a framework for a rational design of this potentially important human therapeutic.

Keywords: amino acid sequence; binding sites; comparative study; crystallography X-ray; fibroblast growth factor.kgf.fgf10: chemistry (ch); human; molecular sequence data; protein conformation; protein folding

KGF is a member of the fibroblast growth factor (FGF) family of proteins and was originally isolated from the conditioned medium of a human embryonic lung fibroblast cell line (Rubin et al., 1989). The FGF family of mitogenic proteins consists of at least 14 polypeptides identified as acidic FGF (aFGF) (Jaye et al., 1986), basic FGF (bFGF) (Abraham et al., 1986), FGF3 or INT-2 (for INTEgration-2) (Dickson & Peters, 1987), FGF4 or HST (for Human Stomach Cancer), or K-FGF (K for Kaposi's sarcoma) (Delli & Basilico, 1987), FGF5 (Zhan et al., 1988), FGF6 (Marics et al., 1989), FGF7 or KGF (Rubin et al., 1989), FGF8 (Tanaka et al., 1992), and FGF9 (Miyamoto et al., 1994), and FGF10 (Yamasaki et al., 1996). The first 10 members of this family share considerable sequence and structural similarity encompassing distinct as well as overlapping biological activities (Burgess & Maciag, 1989;

Gospodarowicz, 1990; Klagsbrun & D'Amore, 1991; Ledoux et al., 1992; Kiefer et al., 1993; Gimenezgallego & Cuevas, 1994; Ohuchi et al., 1994; Rubin et al., 1995). Most recently four new members of the family have been identified as fibroblast growth factor homologous factors (FHF1–4) that share less than 30% identity with the other FGFs (Smallwood et al., 1996). KGF induces its biological activity on keratinocytes and influences proliferation and differentiation patterns of multiple epithelial cell lineages within skin, lung, and the reproductive tract, and therefore, may have potential therapeutic benefits as a mediator in the growth and development of the related tissues. The sequence alignment of KGF against acidic and basic FGF structures indicates that the homologous segment from residue 33 through residue 163 of KGF has an approximately 40% identity with FGF (Finch et al., 1989) (Fig. 5a). The ability of the FGF family of proteins to bind heparin is a distinguishing feature for this family of proteins. However, the role of heparin may be of a dual nature, depending on the growth factor it binds. Heparin binding is considered essential for

Reprint requests to: Timothy D. Osslund, Amgen, Molecular Structure and Design Group, Amgen Center, Thousand Oaks, California, 91320-1789; e-mail: osslund@amgen.com.

Table 1. Truncation mutants of recombinant KGF and their average mitogenic activity^a

KGF N-terminal truncation analog	Avg. % activity
dN15	100
dN16	100
dN17	110
dN18	120
dN19	130
dN21	90
dN22	90
dN23	100
dN24	100
dN25	50
dN26	20
dN27	3

^aMutants were obtained as described in text. Activity was determined based on tritiated thymidine uptake in Balb/MK cells and averaged over three doses of the truncated forms of KGF.

biological activity for bFGF (Moy et al., 1997), whereas heparin binding to KGF has been reported as inhibitory in Balb/MK cells expressing the KGFR (Reich-Slotky et al., 1994). Recent reports have suggested that heparin may be a dual modulator of biological activity (Bonneh-Barkay et al., 1997).

The FGF families of proteins bind to specific cell surface signal transducing receptors corresponding to a subgroup of the tyrosine kinase family. Four distinct FGF receptors (FGFR) have been isolated and identified as FGFR1–FGFR4 (Kornbluth et al., 1988; Ruta et al., 1988, 1989; Keegan et al., 1991; Partanen et al., 1991; Givol & Yayon, 1992; Johnson & Williams, 1993) including several isoforms of FGFR1 and FGFR2 (Johnson et al., 1990; Reid et al., 1990; Champion et al., 1991; Crumley et al., 1991; Eisemann et al., 1991; Miki et al., 1992). Analysis of the amino acid

sequences of the FGF receptor family suggests they share a common structural motif that includes three extra cellular immunoglobulin (IgG) like domains, a transmembrane domain, and a cytoplasmic tyrosine kinase domain. Acidic and basic FGF bind to FGFR1–4, whereas KGF binds with highest affinity to isoforms of FGFR2 known as the KGF receptor (KGFR) (Bottaro et al., 1990; Dionne et al., 1990; Keegan et al., 1991; Miki et al., 1991; Ron et al., 1993; Reich-Slotky et al., 1995). The first KGF receptor clones identified had only two Ig domains but subsequent isoforms having three domains that bound KGF with similar affinities were also isolated (Miki et al., 1991, 1992). The Ig-like domain proximal to the transmembrane domain is the most sequence divergent with respect to the other FGF receptors, and is the carboxy-terminal half of this domain, which imparts KGF-specific binding (Finch et al., 1995; Yayon et al., 1995).

There are two proposed receptor binding sites on bFGF, one of which is a cluster of hydrophobic residues that provides approximately 75% of the binding affinity, and a secondary binding site known as the putative receptor binding loop, which has a 250-fold lower binding affinity (Springer et al., 1994). Domain swapping (Seddon et al., 1995) and FGF/KGF chimeric molecules have also been used to identify the two receptor binding sites (Reich-Slotky et al., 1995).

Results and discussion

The mutagenesis and truncation analysis of the amino terminus of KGF revealed that the first 23 residues of the native KGF polypeptide can be removed without reducing mitogenic activity (Table 1). However, sequential deletion of the next six residues dramatically lowered the biological activity. A truncation mutant with the first 23-amino terminus residues deleted and an arginine to a glutamine mutation at position 144 of the native sequence ($\Delta 23$ R144Q) (Fig. 1) retained its biological activity and had greater thermal stability than the native form (Table 2). This KGF mutant was chosen for crystallographic studies because of its ability to produce high resolution diffraction quality crystals. Crystals of the

Fig. 1. The amino acid sequence of r-hu-KGF. The subject of this study is a mutant of the native polypeptide in which the first 23 residues have been deleted and a substitution of GLN has replaced ARG at position 144. Continued truncation of the protein beyond residue S24 resulted in loss of biological activity (Table 1). Specific residues that have been evaluated using alanine-scanning mutagenesis are gray filled and are the site of the receptor specific binding site.

Table 2. A comparison of the biological activity and thermal stability of different analogs of Δ N23 KGF and alanine scanning of the putative receptor binding loop residues 120–133^a

Analog	Average (%)		Thermal transition point (°C) ^d	Percent solvent accessibility ^e
	Activity (%) ^b	Elisa ^c		
KGF Δ N23	100	1	57	
KGF 23R144Q	100	0.99	63	32
KGF Δ N23 S122A	70	0.56	57	0
KGF Δ N23 K124A	100	0.22	57	35
KGF Δ N23 W125A	1	0.41	57	28
KGF Δ N23 T126A	60	0.61	54	31
KGF Δ N23 H127A	50	0.84	57	32
KGF Δ N23 N128A	100	0.9	57	84
KGF Δ N23 G129A	90	0.95	50	71
KGF Δ N23 G130A	80	1.03	47	43
KGF Δ N23 E131A	70	0.5	43	34
KGF Δ N23 M132A	2	0.34	34 ^f	0
KGF Δ N23 F133A	2	0.46	34 ^f	2

^aThe mutation Δ 23R144Q increased thermal stability of the protein compared to the native truncated form and still retained full biological activity. Residues with little solvent accessibility presumably lose biological activity because of a perturbation in the hydrophobic core of the protein.

^bApproximate average mitogenic signal as percent of KGF Δ N23 standard activity over linear range of standard curve. Tritiated thymidine uptake in Balb/MK cells.

^cELISA with monoclonal anti-KGF antibodies; sample signal over standard signal. (Standard deviation <0.2.)

^dApparent T_m , the midpoint temperature in the thermal transition curve obtained by following changes in CD signal at 235 nm. This is an irreversible reaction, so this temperature does not represent a true thermodynamic parameter ($\pm 1^\circ\text{C}$).

^eSolvent accessible calculations were determined with XPLOD; total accessible surface area of each residue was divided by the total possible surface area of that residue type.

^fSome precipitation at 4°C . Thermal transition could have begun before measurements started.

full-length KGF were obtained but they did not diffract with sufficient resolution ($7.0\text{ \AA}$) to allow further investigation.

Crystal structure of Δ 23R144Q KGF

The three-dimensional structure of Δ 23R144Q KGF was solved at a $1.6\text{ \AA}$ resolution by the multiple isomorphous replacement including the anomalous scattering method (MIRAS) (Table 3). The entire model was built into an electron density map calculated with solvent flattened MIRAS phases obtained from uranyl acetate and potassium tetra nitro-platinate (II) derivatives (Fig. 2). KGF adopts the β -trefoil motif (Murzin et al., 1992) and is similar to the structures of Interleukin 1 alpha (Graves et al., 1990), IL-1 beta (Priestle et al., 1988; Finzel et al., 1989), the Kunitz-type soybean tyrosine inhibitor family (Sweet et al., 1974; Onesti et al., 1991), the bi-functional protease K/ α -amylase inhibitor, aFGF and bFGF (Eriksson et al., 1991; Zhang et al., 1991; Zhu et al., 1991; Blaber et al., 1996) (Fig. 3).

KGF has 10 well-defined beta strands, which form five double-stranded anti-parallel beta sheets (Fig. 4). In the loop between residues 133 and 144, there is a single beta-sheet hydrogen bond between residues 137 and 141, which has been identified in the

Fig. 2. MIR electron density map of the carboxyl terminus of KGF (Δ 23R144Q). The electron density of the MIR solvent flattened map was calculated at $2.7\text{ \AA}$ and contoured at 2.0σ . The refined atomic model of the last five residues in the carboxy terminus of KGF has been superimposed into the density.

crystal FGF structures as a sixth beta-sheet. This poorly defined sixth beta-strand pair has been identified as having a helix-like nature in the high resolution bFGF NMR structure (Moy et al., 1996). Despite these subtle differences the structures of KGF, aFGF, and bFGF are very similar with an RMS deviation (RMSD) of about $0.89\text{ \AA}$ based on common C-alpha atoms.

The most striking difference between the KGF and FGF structures is the ordered N-terminus of the KGF (Δ 23R144Q) mutant. The amino termini of both acidic and basic FGF are disordered presumably as consequence of a *cis/trans* isomerization of the prolines, which are amino proximal to a common basic residue (Lys-12 aFGF, Lys22-bFGF) (Fig. 5A). The KGF structure does not share this proline isomerization, and the first 12 amino acids of the truncated KGF can clearly be placed in electron density. In addition, there are structural elements that contribute to the relative stability of the amino terminus of Δ 23R144Q-R144Q that include a four-residue extension at the amino terminus of the first beta strand, a common glycine turn that causes the first five residues of Δ 23KGF-R144Q to fold back toward the core structure, and a main-chain hydrogen bond formed between the backbone nitrogen of TYR-25 and the carbonyl of CYS-102, which is a part of a nonconserved disulfide bond.

The most significant deviation within the core of the FGF family of proteins is located in the putative receptor binding loop (Baird et al., 1988) between residues 124 and 132 (Fig. 5B) of KGF. Swapping of the analogous loop in aFGF and bFGF changed the

Table 3. Summary of crystallographic data and results^a

	NAT11	NAT12	UA1	PTNP1	UA2	PNT2
Resolution (Å)	2.2	1.5	2.1	2.2	2.1	2.2
Cell parameters (Å)						
A	69.2	69.2	69.2	69.0	69.1	69.0
B	69.0	68.6	68.9	69.2	69.4	69.2
C	66.7	65.9	65.7	66.4	66.3	66.5
Total observations	39,185	60,197	40,825	25,726	67,510	61,239
Unique reflections	8,353	21,552	9,376	7,951	9,462	8,039
Completeness (%)	98.6 (98.5)	84.6 (27.0)	93.4 (90.5)	93.7 (92.0)	92.8 (88.0)	94.6 (92.4)
R_{sym} (%) ^b	4.2 (8.0)	3.9 (31.3)	4.6 (12.3)	2.4 (5.6)	4.9 (18.0)	3.8 (9.4)
Soaking conditions						
Concentration (mM)			1.0	0.1	1.0	0.1
Time (h)			24	24	24	24
Number of sites			2	2	2	2
R_{iso} (%) ^c			13.1	10.6	10.8	9.0
R_c (%) ^d			63.0	56.0	68.0/38.0	55.0/37.0
Phasing power ^e			1.2	2.4	1.4/1.8	2.4/2.0
Resolution range for refinement (Å), 2σ cutoff		6.0–1.5				
Total reflections in the refinement		19,819				
$R_{\text{cryst}}/R_{\text{free}}$ (%) ^f		21.7/27.9				
No. waters		92				
RMSD on bond length (Å)		0.013				
RMSD on bond angle (°)		1.768				
Average B-factor (Å ²)		26.6				

^aValues in parentheses are for highest resolution shell.

^b $R_{\text{sym}} = \sum |I_{\text{obs}} - I_{\text{ave}}| / \sum I_{\text{ave}}$, unweighted R -factor on I among symmetry mates.

^c $R_{\text{iso}} = \sum ||F_{\text{ph}}| - |F_p|| / \sum |F_p|$, where F_p and F_{ph} are the structure factor amplitudes of the NAT11 and derivative data, respectively.

^d $R_c = \sum ||F_{\text{ph}}| \pm |F_p| - |F_h| / \sum ||F_{\text{ph}}| \pm |F_p||$ for centric reflections. The second entry is for anomalous dataset where $R_c = \sum ||F_{\text{ph}}| - |F_{\text{ph}}^+|| + ||F_{\text{ph}}| - |F_{\text{ph}}^-|| / \sum ||F_{\text{ph}}^+| + |F_{\text{ph}}^-||$ calculated for the top 25% largest bijvoet differences. Data cutoff >2.5 Å and $F/\sigma = 4.0$.

^ePhasing power is RMS ($|F_h|/E$), where F_h is the calculated heavy atom structure factor amplitude, and E is the residual lack of closure error. The second entry is for anomalous(ano) phasing power, RMS ($|F_h'|/E_{\text{ano}}$), where F_h' is the anomalous correction component amplitude. Data cutoff >2.5 Å and $F/\sigma = 4.0$.

^f $R_{\text{cryst}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$, where F is the structure factor. R_{free} is the cross-validation R -factor computed for the test set of reflections (10% of the total), which are omitted from the refinement (Brüger, 1992).

specificity in receptor binding (Seddon et al., 1995). Mutational analysis of this loop in bFGF (residues 107–116) indicates that LYS 111, TYR-112, and TRP-115 are the essential residues for a second receptor binding site, which confers distinct biological activity (Springer et al., 1994). The conserved LYS at position 124 of KGF (LYS 101 aFGF, LYS 111 bFGF) is the point at which the ligand-specific loop begins its largest difference in the FGF family and regains its structural similarity at Met 132 (TRP 107 aFGF, TRP 115 bFGF). The additional of two residues in the case of aFGF and four residues in the case of KGF with respect to bFGF causes the difference in the structure of this loop in the FGF protein family. The hydrogen bonding between residues His 127 and GLY 130 provide the structural stability of a beta-hairpin turn for the KGF loop. When the loop is superimposed (Fig. 5B), LYS 111, Tyr-112, and TRP-115 of bFGF structurally correspond to residues LYS 124, TRP 125, and Met 132 of KGF.

Springer et al. demonstrated that alanine mutagenesis of the conserved lysine reduced biological activity; however, this does not appear to be the case for KGF (Table 2). Our results indicate that alanine mutagenesis of the conserved lysine of KGF does not effect biological activity; however, the next three residues are essential. These residues (TRP 125, THR 126, and His 127) struc-

turally form an exposed face on one side of the loop. Notably, the mutation of the most distal and exposed residue ASN 128 of the KGF loop does not appear to affect biological activity.

Mutagenesis of the structurally related TYR 112 (bFGF) and TRP 125 (KGF) both result in lower biological activity (Table 2). Elisa and stability data suggest that the lower biological activity of alanine mutagenesis of Met 132 and also PHE 133 is caused by a disruption of the hydrophobic core of KGF and not specific interaction of these residues with the receptor. A correlation between the mutagenesis, biophysical, and structural data suggestions that receptor binding specificity is principally conferred by residues TRP 125, THR 126, and His 127 of the KGF molecule.

Heparin binding site

Superimposition of the KGF structure on the bFGF structure with bound heparin (Faham et al., 1996) suggests residues that may be responsible for heparin binding in KGF (Fig. 6A). The superimposition reveals that only GLN-152 of KGF and GLN-135 of FGF are identical in the heparin-binding pocket (Fig. 6B). However, the hydroxyl group of THR-42 of KGF can potentially replace the high affinity binding of the carbonyl of Asn-28 of bFGF. In the MIRAS

Fig. 3. Overlay of the mutant KGF ($\Delta 23R144Q$) with the aFGF and bFGF. The superimposition of the atoms was performed within the Protein Design module of Quanta (Molecular Simulation Inc., 1997). The amino sequences of each protein were initially aligned and the structures were superimposed using the amino acids, which matched identically. The RMSD between the three structures when calculated based on the C alpha positions was 0.89 Å. The KGF mutant ($\Delta 23R144Q$) is colored red, and the aFGF and bFGF structures are colored green and blue. The PDB I.D. codes are 2AFG and 1BAS for aFGF and bFGF, respectively.

electron density map of KGF, it is clear that the side chain of GLN 138 has two distinct conformations. When the KGF and bFGF heparin structures are superimposed, the GLN-138 could reasonably bind to either of two positions on heparin. The nonpolar residue Val-143 of KGF superimposes with the high-affinity residue Lys-126 of bFGF, and thus cannot participate in heparin binding. However, when the structures are superimposed, residue THR-154 of KGF appears to be pointing toward the heparin-binding pocket, and thus may contribute the binding energy lost by the valine.

Thermal stability of mutant

Mutational analysis around the heparin binding site of KGF serendipitously provided a mutant of KGF with increased thermal stability. The crystallographic structure of this mutant indicates the glutamine side chain interacts with the backbone atoms of an adjacent loop (Fig. 7). An increase in the thermal stability of this mutant is coincident with the formation of two new hydrogen bonds that could not have been made with the native arginine side chain. Additional salt bridges identified in this region of the structure may be formed in the native structure, but clearly would be favored in the thermally stable mutant.

Implication of the structure

The resolution of the structure of a third member of the FGF family of proteins reinforces the relative conserved structure and

topology of this biologically important class of proteins. Based on the structural similarities of these proteins, the biological specificity must be conferred by specific differences in amino acids and not global structural changes. An additional four residues of the putative receptor specific binding site apparently provides the structural framework for distinction between members of the FGF family.

Because the biological activity is essentially similar between the native and truncated forms of KGF, it is suggested that the first 23 residues play a limited structural and functional role. The next 10 residues of the $\Delta 23R144Q$ KGF are well ordered, but not in the case for the other FGF structures. It is noteworthy to point out that $\Delta 23R144Q$ KGF begins at the residue identified as the last residue that can be deleted before losing its biological activity. This may suggest that the hydrogen bond between this residue and the non-conserved disulfide bond is an essential structural element required for biological activity.

Despite the similarity of structure, the difference in the biological activity of the FGF family of proteins when binding to heparin may be correlated to the amino acid differences within their respective heparin binding pockets. Four of the seven residues identified in the bFGF/heparin structure are similar to those of KGF. A notable exception to the conserved heparin-binding motif, is a Valine at position 143, which structurally corresponds to a high-affinity lysine in the bFGF/heparin structure.

The rational design of thermally stable protein mutants remains poorly understood (Querol et al., 1996). Identification of novel

Fig. 4. KGF secondary structure. The KGF structure has 10 well-defined beta strands that form five double-stranded anti-parallel beta sheets. In addition, the beta strand $\beta 4$ is the center strand in a triple stranded beta sheet. This triple-stranded sheet is unique between KGF and the FGF structures, and is created because beta strand $\beta 1$ is significantly longer than its counterparts in either acidic or basic FGF. A sixth poorly defined β -strand pair is in the loop between residues 133 and 144, and is defined by only a single hydrogen bond between the two strands.

protein analogous with improved pharmacological properties that retain biological activity is generally empirically developed. The correlation between the thermal stability, heparin binding, and biology of other members of the FGF family continues to provide invaluable information in the development of KGF as a human therapeutic.

Experimental procedures

Production of KGF deletion and substituted mutants

Oligonucleotide linkers were designed to generate the desired amino-terminal deletion for each mutant with convenient restriction endonuclease sites at either end. The 5' end of each linker

contained the desired N-terminal truncated KGF sequence. The 3' end was constant for all constructs in the series, containing a Mlu I site in the KGF sequence 3' of the N-terminus. Each KGF truncation mutant protein was purified from *Escherichia coli* cultures grown and induced under conditions that optimize expression of recombinant protein. Soluble KGF protein was purified over Sep-Pak® Vac 6cc Accell™ Plus CM cartridges (Waters division of Millipore). After washing with 0.1 M NaCl to remove the majority of *E. coli* contaminant proteins, KGF proteins were eluted in 0.6 M NaCl, 20 mM Na₂PO₄, pH 7.5.

Measurement of thermostability of KGF mutants

The CD spectra were determined on a Jasco J-720 spectropolarimeter controlled by a DOS-based computer and Jasco software

Fig. 5. A: The sequence alignment between the mutant KGF (R144Q) and the acidic and basic forms of FGF. The sequences were obtained from the acid and basic FGF PDB coordinates, 2AFG and 1BAS, respectively. Gaps between the alignments are indicated by spaces and residues, amino acids that are identical in the alignment are indicated with a line between sequences (Quanta97, Molecular Simulations Inc., 1997). The receptor specificity loop is bold. Unless otherwise noted, amino acid numbering referenced within the text has been adjusted to reflect this figure. **B:** The putative receptor-specific binding site for KGF. The structures of bFGF and the mutant KGF ($\Delta 23R144Q$) were superimposed as described in Figure 3. This loop, which is composed of residues 120–133 of KGF, is suggested to be the receptor-binding site, which confers the biological specificity between the FGF family of proteins. $\Delta 23R144Q$ KGF is depicted in red, and bFGF is blue. Residues of KGF, which may be contributing to receptor binding, are labeled, respectively.

using cylindrical cuvettes with a path length of 0.02 cm for the far-UV regions (235 nm). Thermal denaturation was performed on the same instrument using a Peltier thermal control JTC-343 unit and a rectangular cuvette having a 0.1 cm path length with protein concentration of 0.2 mg/mL. The thermal stability temperature (Table 2) is that temperature in which the CD spectra at 235 nm began to change ($\pm 1^\circ\text{C}$).

The thermal-induced denaturation of KGF is an irreversible reaction, but is extremely reproducible under the conditions em-

ployed here, and can be used to indicate the temperature above which the proteins no longer remain folded and soluble (Chen et al., 1994).

Crystallization, data collection, phasing, model building, and refinement

KGF crystals were grown by vapor diffusion from a crystal growth solution (100 mM sodium acetate pH 4.7, 14.7% polyeth-

Fig. 6. A: Potential heparin binding site. The structure of bFGF complexed with heparin (PDB identification code 1BFB) was superimposed on the $\Delta 23R144Q$ KGF structure (Quanta97, Molecular Simulations Inc., 1997). These coordinates were imported into Web Lab Viewer (Molecular Simulations Inc., 1997), and a transparent surface representation of the heparin structure was created. The amino acids of KGF, which are implicated in heparin binding, are shown along with an alpha carbon ribbon tracing. Carbon atoms are colored white, nitrogen blue, and oxygen red. **B:** Alignment of heparin binding residues of bFGF with KGF. Residues of bFGF, which bind to heparin and the corresponding residues of KGF based on structural alignment. bFGF residues with side chains that interact with heparin with high affinity are marked with an asterisk.

ylene glycol methyl ether (PME) 2 K, 13.0 mM ammonium sulfate) at a protein concentration of 45.0 mg/mL, using a modified seeding technique. The crystals grow in I222 space group, $a = 69.2$, $b = 68.6$, $c = 65.9$ Å and diffract to 1.6 Å. All diffraction data were collected on an R-Axis Ilc imaging system installed on an 18 kW Rigaku RU300 generator using a single crystal for each data set. These data were processed with the Molecular Structure Corporation (MSC) data processing software, version 2.0, as well as with the HKL software package (Otwinowski & Minor, 1996). To produce sufficient data collection quality crystals for a rapid evaluation of the heavy atoms derivatives, 20 crystallization trays (Linbro Plates, Flow Laboratories, McLean, Virginia), utilizing 16 wells per plate with a pre-tested crystal growth solution, 10-mL drop sizes, and 18 °C temperature, were set up using an ICN crystallization robot. This method produced approximately 160 wells each containing 10 to 15 diffraction quality crystals (approximately $0.3 \times 0.5 \times 0.7$ mm). Heavy atom compounds were screened for their appropriate concentrations and the soaking times by selecting five drops with diffraction quality crystals for each compound and setting them up at 0.1, 0.5, 1.0, 2.0, 4.0, and >25.0 mM concentrations. Crystals were checked for their integrity and diffraction quality starting from the highest concentration of the heavy atom

compound at 3 h, 1 day, 3 day, 1 week, and 1 month time intervals. If a crystal was suitable (diffracting better than 3.0 Å), a data set was collected and evaluated for its isomorphousness and the extent of derivatization using XTALVIEW (Molecular Simulations Inc., 1997) and PHASES (Furey & Swaminathan, 1997). Heavy atom sites were identified in difference Patterson and difference Fourier maps, refined for their positions and occupancies, and confirmed by examining the peaks in cross-phased difference Fourier maps. Out of 19 derivative data sets collected, two derivatives, uranyl acetate (UA1) and potassium tetra nitro platinate (II), (PTNP1) were deemed most useful based on their phasing powers, figures of merit, and Rculis (Table 1). The MIR map calculated with phases obtained from the isomorphous differences of UA1 and PTNP1 derivatives with the native (NAT11) data showed clear molecular boundaries and large regions of continuous protein like density. The quality of MIR electron density map improved (figure-of-merit increase from 0.539 to 0.773 at 2.5 Å) by solvent flattening (with 54% bulk solvent) using PHASES. Careful inspection of this map revealed sufficient main-chain and side-chain density for an unambiguous identification of the sequence of the 12 carboxyl terminus residues of KGF (Fig. 2) and established a well-defined landmark region to assist in tracing of the map. Model building

Fig. 7. Analysis of the mutation of ARG 144 to GLN. The mutation of $\Delta 23$ KGF ARG144 to GLN increased the thermal stability of the protein from 49 to 54 °C (Table 2). The increased stability of the mutant is caused by a network of hydrogen bonds formed between the N_ϵ of GLN144 and the carbonyl O of GLU-93 and between the O_ϵ of GLN144 and the backbone N of LYS146. Additional formation of salt bridges between the GLU-93 side chain with LYS 92 and 146 also contributes to the overall structural stability of the protein. Hydrogen bonds are indicated as a yellow line and distances measured in angstroms are noted.

was done using the X-FIT module of QUANTA (Molecular Simulations Inc., 1997). Initially, seven long segments containing a total of 84 residues were built into the density. To further enhance the quality of the map, uranyl acetate (UA2) and potassium tetra nitro platinate (II), (PTNP2) derivative datasets were recollected for the completeness of anomalous data. A solvent-flattened MIRAS map, incorporating isomorphous differences from UA1, PTNP1, UA2, and PTNP2 and anomalous differences from UA2 and PTNP2 looked much improved (figure of merit 0.859 at 2.5 Å) to the previously available MIR map and allowed the connection of the discontinuous segments into a single chain. The initial model with 139 residues (8 residues with no side chains built) and 24 water molecules yielded a crystallographic R -factor of 0.196 (8.0–2.5 Å, $I/\sigma(I) > 2$) with the use of simulated annealing, positional, and temperature factor refinements using XPLOR (Brünger, 1992). At this stage a new native dataset (NAT12) to 1.5 Å was collected and introduced into refinement and further model building. Several cycles of model building, refinement, and phase combination at 1.5 Å resolution resulted into a complete model including four side chains with alternate conformations, 91 solvent molecules, and crystallographic $R_{\text{conventional}}/R_{\text{free}}$ of 0.217/0.279 (6.0–1.5 Å, $I/\sigma(I) > 2$).

Acknowledgments

The first two authors contributed equally to the crystallographic analysis.

References

- Abraham JA, Whang JL, Tumolo A, Mergia A, Friedman J, Gospodarowicz D, Fiddes JC. 1986. Human basic fibroblast growth factor: Nucleotide sequence and genomic organization. *EMBO J* 5:2523–2528.
- Baird A, Schubert D, Ling N, Guillemin R. 1988. Receptor- and heparin-binding domains of basic fibroblast growth factor. *Proc Natl Acad Sci USA* 85:2324–2328.
- Blaber M, DiSalvo J, Thomas KA. 1996. X-ray crystal structure of human acidic fibroblast growth factor. *Biochemistry* 35:2086–2094.
- Bonneh-Barkay D, Shlissel M, Berman B, Shaoul E, Admon A, Vlodavsky I, Carey DJ, Asundi VK, Reich-Slotky R, Ron D. 1997. Identification of glypican as a dual modulator of the biological activity of fibroblast growth factors. *J Biol Chem* 272:12415–12421.
- Bottaro DP, Rubin JS, Ron D, Finch PW, Florio C, Aaronson SA. 1990. Characterization of the receptor for keratinocyte growth factor. Evidence for multiple fibroblast growth factor receptors. *J Biol Chem* 265:12767–12770.
- Brünger AT. 1992. *X-PLOR, Version 3.1: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University.
- Burgess WH, Maciag T. 1989. The heparin-binding (fibroblast) growth factor family of proteins. *Annu Rev Biochem* 58:575–606.
- Champion A-P, Ronsin C, Gilbert E, Gesnel MC, Houssaint E, Breathnach R. 1991. Multiple mRNAs code for proteins related to the BEK fibroblast growth factor receptor. *Oncogene* 6:979–987.
- Chen BL, Arakawa T, Hsu E, Narhi LO, Tressel TJ, Chien SL. 1994. Strategies to suppress aggregation of recombinant keratinocyte growth factor during liquid formulation development. *J Pharmaceut Sci* 83:1657–1661.
- Crumley G, Bellot F, Kaplow JM, Schlessinger J, Jaye M, Dionne CA. 1991. High-affinity binding and activation of a truncated FGF receptor by both aFGF and bFGF. *Oncogene* 6:2255–2262.
- Delli B-P, Basilico C. 1987. Isolation of a rearranged human transforming gene following transfection of Kaposi sarcoma DNA. *Proc Natl Acad Sci USA* 84:5660–5664.
- Dickson C, Peters G. 1987. Potential oncogene product related to growth factors [letter]. *Nature* 326:833.
- Dionne CA, Crumley G, Bellot F, Kaplow JM, Searfoss G, Ruta M, Burgess WH, Jaye M, Schlessinger J. 1990. Cloning and expression of two distinct high-affinity receptors cross-reacting with acidic and basic fibroblast growth factors. *EMBO J* 9:2685–2692.
- Eisemann A, Ahn JA, Graziani G, Tronick SR, Ron D. 1991. Alternative splicing generates at least five different isoforms of the human basic-FGF receptor [published erratum appears in *Oncogene* 6:2379; 1991]. *Oncogene* 6:1195–1202.
- Eriksson AE, Cousens LS, Weaver LH, Matthews BW. 1991. Three-dimensional structure of human basic fibroblast growth factor. *Proc Natl Acad Sci USA* 88:3441–3445.
- Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC. 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science* 271:1116–1120.
- Finch PW, Cunha GR, Rubin JS, Wong J, Ron D. 1995. Pattern of keratinocyte growth factor and keratinocyte growth factor receptor expression during mouse fetal development suggests a role in mediating morphogenetic mesenchymal-epithelial interactions. *Dev Dynam* 203:223–240.
- Finch PW, Rubin JS, Miki T, Ron D, Aaronson SA. 1989. Human KGF is FGF-related with properties of a paracrine effector of epithelial cell growth. *Science* 245:752–755.
- Finzel BC, Clancy LL, Holland DR, Muchmore SW, Watenpaugh KD, Einspahr HM. 1989. Crystal structure of recombinant human interleukin-1 beta at 2.0 Å resolution. *J Mol Biol* 209:779–791.
- Furey W, Swaminathan S. 1997. *PHASES-95: A program package for processing and analyzing diffraction data from macromolecules*. San Diego: Academic Press Inc.
- Gimenezgalligo G, Cuevas P. 1994. Fibroblast growth-factors, proteins with a broad-spectrum of biological-activities. *Neurol Res* 16:313–316.
- Givol D, Yayon A. 1992. Complexity of FGF receptors: Genetic basis for structural diversity and functional specificity. *FASEB J* 6:3362–3369.
- Gospodarowicz D. 1990. Fibroblast growth factor. Chemical structure and biologic function. *Clin Orthopaed Related Res N* 257:231–248.
- Graves BJ, Hatada MH, Hendrickson WA, Miller JK, Madison VS, Satow Y. 1990. Structure of interleukin 1 alpha at 2.7-Å resolution. *Biochemistry* 29:2679–2684.
- Jaye M, Howk R, Burgess W, Ricca GA, Chiu IM, Ravera MW, O'Brien SJ, Modi WS, Maciag T, Drohan WN. 1986. Human endothelial cell growth factor: Cloning, nucleotide sequence, and chromosome localization. *Science* 233:541–545.
- Johnson DE, Lee PL, Lu J, Williams LT. 1990. Diverse forms of a receptor for acidic and basic fibroblast growth factors. *Mol Cell Biol* 10:4728–4736.
- Johnson DE, Williams LT. 1993. Structural and functional diversity in the FGF receptor multigene family. *Adv Cancer Res* 60:1–41.
- Keegan K, Johnson DE, Williams LT, Hayman MJ. 1991. Isolation of an additional member of the fibroblast growth factor receptor family, FGFR-3. *Proc Natl Acad Sci USA* 88:1095–1099.
- Kiefer P, Mathieu M, Close MJ, Peters G, Dickson C. 1993. FGF3 from *Xenopus laevis*. *EMBO J* 12:4159–4168.
- Klagsbrun M, D'Amore PA. 1991. Regulators of angiogenesis. *Annu Rev Physiol* 53:217–239.
- Kornbluth S, Paulson KE, Hanafusa H. 1988. Novel tyrosine kinase identified

- by phosphotyrosine antibody screening of cDNA libraries. *Mol Cell Biol* 8:5541–5544.
- Ledoux D, Gannoun Z-L, Barritault D. 1992. Interactions of FGFs with target cells. *Prog Growth Factor Res* 4:107–120.
- Marics I, Adelaide J, Raybaud F, Mattei MG, Coulier F, Planche J de L-O, Birnbaum D. 1989. Characterization of the HST-related FGF.6 gene, a new member of the fibroblast growth factor gene family. *Oncogene* 4:335–340.
- Miki T, Bottaro DP, Fleming TP, Smith CL, Burgess WH, Chan AM, Aaronson SA. 1992. Determination of ligand-binding specificity by alternative splicing: Two distinct growth factor receptors encoded by a single gene. *Proc Natl Acad Sci USA* 89:246–250.
- Miki T, Fleming TP, Bottaro DP, Rubin JS, Ron D, Aaronson SA. 1991. Expression cDNA cloning of the KGF receptor by creation of a transforming autocrine loop. *Science* 251:72–75.
- Miyamoto T, Takeuchi A, Hayashi H, Onozaki K. 1994. Novel growth promoting activity with a wide target cell spectrum is present in extracts of various cell types. *Biochem Mol Biol Int* 32:973–981.
- Moy FJ, Safran M, Seddon AP, Kitchen D, Boehlen P, Aviezer D, Yayon A, Powers R. 1997. Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor. *Biochemistry* 36:4782–4791.
- Moy FJ, Seddon AP, Boehlen P, Powers R. 1996. High-resolution solution structure of basic fibroblast growth factor determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry* 35:13552–13561.
- Murzin AG, Lesk AM, Chothia C. 1992. Beta-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors. *J Mol Biol* 223:531–543.
- Ohuchi H, Yoshioka H, Tanaka A, Kawakami Y, Nohno T, Noji S. 1994. Involvement of androgen-induced growth factor (FGF-8) gene in mouse embryogenesis and morphogenesis. *Biochem Biophys Res Commun* 204:882–888.
- Onesti S, Brick P, Blow DM. 1991. Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds. *J Mol Biol* 217:153–176.
- Otwinowski Z, Minor W. 1996. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276.
- Partanen J, Maekela TP, Eerola E, Korhonen J, Hirvonen H, Claesson W-L, Alitalo K. 1991. FGFR-4, a novel acidic fibroblast growth factor receptor with a distinct expression pattern. *EMBO J* 10:1347–1354.
- Priestle JP, Schaer HP, Gruetter MG. 1988. Crystal structure of the cytokine interleukin-1 beta. *EMBO J* 7:339–343.
- Querol E, Perez Pons JA, Mozo Villarias A. 1996. Analysis of protein conformational characteristics related to thermostability. *Protein Eng* 9:265–271.
- Reich-Slotky R, Shaoul E, Berman B, Graziani G, Ron D. 1995. Chimeric molecules between keratinocyte growth factor and basic fibroblast growth factor define domains that confer receptor binding specificities. *J Biol Chem* 270:29813–29818.
- Reich-Slotky R, Bonneh-Barkay D, Shaoul E, Bluma B, Svahn CM, Ron D. 1994. Differential effect of cell-associated heparan sulfates on the binding of keratinocyte growth-factor (Kgf) and acidic fibroblast growth-factor to the Kgf receptor. *J Biol Chem* 269:32279–32285.
- Reid HH, Wilks AF, Bernard O. 1990. Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed in the developing mouse brain. *Proc Natl Acad Sci USA* 87:1596–1600.
- Ron D, Reich R, Chedid M, Lengel C, Cohen OE, Chan AM, Neufeld G, Miki T, Tronick SR. 1993. Fibroblast growth factor receptor 4 is a high affinity receptor for both acidic and basic fibroblast growth factor but not for keratinocyte growth factor. *J Biol Chem* 268:5388–5394.
- Rubin JS, Bottaro DP, Chedid M, Miki T, Ron D, Cunha GR, Finch PW. 1995. Keratinocyte growth factor as a cytokine that mediates mesenchymal-epithelial interaction. *Exs* 74:191–214.
- Rubin JS, Osada H, Finch PW, Taylor WG, Rudikoff S, Aaronson SA. 1989. Purification and characterization of a newly identified growth factor specific for epithelial cells. *Proc Natl Acad Sci USA* 86:802–806.
- Ruta M, Burgess W, Givol D, Epstein J, Neiger N, Kaplow J, Crumley G, Dionne C, Jaye M, Schlessinger J. 1989. Receptor for acidic fibroblast growth factor is related to the tyrosine kinase encoded by the *fms*-like gene (FLG). *Proc Natl Acad Sci USA* 86:8722–8726.
- Ruta M, Howk R, Ricca G, Drohan W, Zabelshansky M, Laureys G, Barton DE, Flanck U, Schlessinger J, Givol D. 1988. A novel protein tyrosine kinase gene whose expression is modulated during endothelial cell differentiation. *Oncogene* 3:9–15.
- Seddon AP, Aviezer D, Li LY, Boehlen P, Yayon A. 1995. Engineering of fibroblast growth factor: Alteration of receptor binding specificity. *Biochemistry* 34:731–736.
- Smallwood PM, Munoz Sanjuan I, Tong P, Macke JP, Hendry SH, Gilbert DJ, Copeland NG, Jenkins NA, Nathans J. 1996. Fibroblast growth factor (FGF) homologous factors: New members of the FGF family implicated in nervous system development. *Proc Natl Acad Sci USA* 93:9850–9857.
- Springer BA, Pantoliano MW, Barbera FA, Gunyuzlu PL, Thompson LD, Herblin WF, Rosenfeld SA, Book GW. 1994. Identification and concerted function of two receptor binding surfaces on basic fibroblast growth factor required for mitogenesis. *J Biol Chem* 269:26879–26884.
- Sweet RM, Wright HT, Janin J, Chothia CH, Blow DM. 1974. Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 angstrom resolution. *Biochemistry* 13:4212–4228.
- Tanaka A, Miyamoto K, Minamino N, Takeda M, Sato B, Matsuo H, Matsumoto K. 1992. Cloning and characterization of an androgen-induced growth factor essential for the androgen-dependent growth of mouse mammary carcinoma cells. *Proc Natl Acad Sci USA* 89:8928–8932.
- Yamasaki M, Miyake A, Tagashira S, Itoh N. 1996. Structure and expression of the rat mRNA encoding a novel member of the fibroblast growth factor family. *J Biol Chem* 271:15918–15921.
- Yayon A, Gray TE, Zimmer Y, Eisenstein M, Givol D. 1995. Ligand-binding by fibroblast growth-factor receptors investigated using chimeric receptor molecules. *J Recept Signal Transduct Res* 15:185–197.
- Zhan X, Bates B, Hu XG, Goldfarb M. 1988. The human FGF-5 oncogene encodes a novel protein related to fibroblast growth factors. *Mol Cell Biol* 8:3487–3495.
- Zhang JD, Cousens LS, Barr PJ, Sprang SR. 1991. Three-dimensional structure of human basic fibroblast growth factor, a structural homolog of interleukin 1 beta [published erratum appears in *Proc Natl Acad Sci USA*. 1991. 88:5477]. *Proc Natl Acad Sci USA* 88:3446–3450.
- Zhu X, Komiya H, Chirino A, Faham S, Fox GM, Arakawa T, Hsu BT, Rees DC. 1991. Three-dimensional structures of acidic and basic fibroblast growth factors. *Science* 251:90–93.