An Atomic Resolution Structure for Human Fibroblast Growth Factor 1

Matthew J. Bernett, Thayumanasamy Somasundaram, and Michael Blaber*
Institute of Molecular Biophysics and Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida

ABSTRACT **A 1.10-Å atomic resolution X-ray structure of human fibroblast growth factor 1 (FGF-1), a member of the β-trefoil superfold, has been determined. The β-trefoil is one of 10 fundamental protein superfolds and is the only superfold to exhibit 3-fold structural symmetry (comprising 3 "trefoil" units). The quality of the diffraction data permits unambiguous assignment of Asn, Gln, and His rotamers, Pro ring pucker, as well as refinement of atomic anisotropic displacement parameters (ADPs). The FGF-1 structure exhibits numerous core-packing defects, detectable using a 1.0-Å probe radius. In addition to contributing to the relatively low thermal stability of FGF-1, these defects may also permit domain motions within the structure. The availability of refined ADPs allows a translation/libration/screw (TLS) analysis of putative rigid body domains. The TLS analysis shows that β-strands 6–12 together form a rigid body, and there is a clear demarcation in TLS motions between the adjacent carboxyl- and amino-termini. Although separate from β-strands 6–12, the individual β-strands 1–5 do not exhibit correlated motions; thus, this region appears to be comparatively flexible. The heparin-binding contacts of FGF-1 are located within β-strands 6–12; conversely, a significant portion of the receptor-binding contacts are located within β-strands 1–5. Thus, the observed rigid body motion in FGF-1 appears related to the ligand-binding functionalities. Proteins 2004;57:626–634.**
© 2004 Wiley-Liss, Inc.

Key words: atomic-resolution; anisotropic displacement parameters; translation/libration/screw tensors; X-ray crystallography; fibroblast growth factor; β-trefoil; protein dynamics

INTRODUCTION

Human acidic fibroblast growth factor (FGF-1), the broadest-specificity human mitogen known, is a member of the β-trefoil superfold and exhibits a characteristic 3-fold structural symmetry (Fig. 1). Proteins that share this architecture include the fibroblast growth factors,[1] the Kunitz soybean trypsin inhibitors,[2] the ricin-like plant and bacterial toxins,[3–6] interleukin-1α and -1β,[7] and the hisactophilin-like actin-bundling proteins,[8] and various carbohydrate-binding proteins such as xylanase.[9] Although these proteins exhibit diverse functionalities, they all share a role as high-affinity ligands.

The β-trefoil fold was first observed in the soybean trypsin inhibitor by Sweet et al.,[2] and further described and characterized by McLachlan[10] and Chothia and coworkers.[11] The structural architecture consists of 12 β-strands that form 6 β-hairpins. The fold can be conceptually divided into a 6-stranded β-barrel "top" half (formed by β-strands 1, 4, 5, 8, 9, and 12) and a β-hairpin triplet "bottom" half (formed by β-strands 2, 3, 6, 7, 10, and 11) that closes off the "bottom" of the β-barrel (Fig. 1). An axis of 3-fold symmetry runs through the center of the β-barrel, dividing the structure into 3 "trefoil" subdomains of approximately 40 amino acids each.

The structure of β-trefoil proteins has been postulated to have arisen from a series of gene duplication events involving the trefoil subdomain.[12–14] However, the trefoil subdomains are not compact, independent regions, and the structural symmetry, combined with a circularly permutable architecture, provides for possible alternative "domain-swapped" definitions for the repeating structural motif. The backbone atoms of each structurally conserved region of the symmetry-related subdomains in FGF-1 overlay with a root-mean-square deviation (RMSD) between 0.6 and 0.8 Å.[15,16] The backbone atoms of the entire molecule can be rotated by the 3-fold symmetry and overlaid upon itself, with an RMSD of 1.1 Å. However, a comparison of the primary structures of the 3 trefoil subdomains suggests a level of identity that is only marginally above random, indicating extensive divergence of the primary structure of the subdomains despite the pronounced tertiary structure symmetry. Specific regions within FGF-1 have been identified as conferring receptor-binding and

The Supplementary Materials referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/index.html

Grant sponsor: Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science; Contract number: W-31-109-Eng-38. Grant sponsor: Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources; Grant number: RR07707. Grant sponsor: National Science Foundation; Grant number: MCB 0314740.

*Correspondence to: Michael Blaber, 406 Kasha Laboratory, Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4380. E-mail: blaber@sb.fsu.edu

Received 2 March 2004; Accepted 20 May 2004

Published online 29 July 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20239

© 2004 WILEY-LISS, INC.

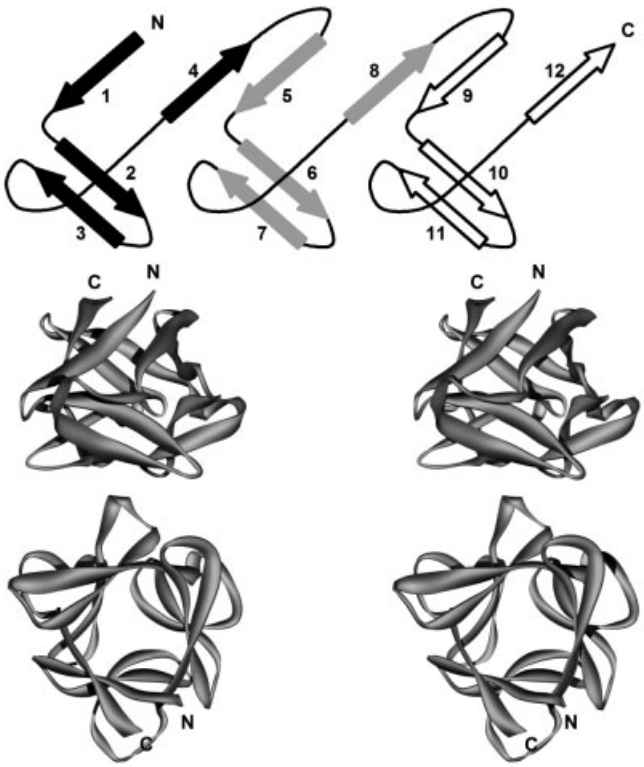

Fig. 1. **Top panel**: A schematic representation of the secondary structure of FGF-1, and the numbering of the individual β-strands. The β-strands comprising the "trefoil" subdomains 1–3 are indicated by the black, gray, and white shading, respectively. The amino and carboxyl termini are indicated. **Middle panel**: A relaxed stereo ribbon diagram of FGF-1 as a "side" view. **Bottom panel**: A relaxed stereo diagram looking down the 3-fold axis of structural symmetry.

TABLE I. Crystal, Data Collection, and Refinement Statistics for FGF-1

Crystal	
Space group	$C\,2\,2\,2_1$
Crystal cell constants	a = 74.26 Å, b = 96.21 Å, c = 108.96 Å
Resolution (Å)	24.71–1.10
Mosaicity (°)	0.30
Data collection	
Reflections measured	1,925,881
Reflections unique	151,310
Redundancy	12.7
I/σ, overall	72.7
I/σ, (1.13–1.10 Å)	4.1
R_{merge} (%), overall	6.1
R_{merge} (%), (1.13–1.10 Å)	38.2
Completeness (%), overall	96.4
Completeness (%), (1.13–1.10 Å)	70.7
Refinement	
No. of nonhydrogen protein atoms	2294
No. of solvent molecules	407
R_{cryst} (%)	14.7
R_{free} (%)	17.1
RMS bond length deviations (Å)	0.015
RMS bond angle deviations (°)	2.285
Mean B-factor, protein (Å^2)	19.4
Mean B-factor, solvent (Å^2)	32.7
Ramachandran most favored (%)	93.9
Ramachandran additional allowed (%)	5.7
Ramachandran generously allowed (%)	0.4
Ramachandran disallowed region (%)	0.0

heparin-binding functionalities,[17–19] and these contribute to the divergence of the subdomains.

In this report, we describe the details of an atomic resolution (1.10 Å) structure of human FGF-1, including characterization of core-packing defects within the structure. Refinement of atomic anisotropic displacement parameters (ADPs), and subsequent evaluation of translation/libration/screw (TLS) motions within the FGF-1 structure, identifies β-strands 6–12 as comprising a rigid body domain, while β-strands 1–5 appear relatively flexible. The rigid body demarcation of these regions follows the general segregation of receptor-binding and heparin-binding functionalities within FGF-1.

MATERIALS AND METHODS

Data Collection and Processing

The 140–amino acid form of human FGF-1, containing an additional amino-terminal 6-residue His-tag, was expressed and purified as previously described.[20] The purified protein was dialyzed against 50 mM HNa_2PO_4, 100 mM NaCl, 10 mM $(NH_4)_2SO_4$, 2 mM dithiothreitol (DTT), and 0.5 mM ethylenediaminetetraacetic acid (EDTA), pH 5.8, for crystallization. Crystals were grown using the hanging-drop vapor-diffusion method with a protein concentration of 12 mg/mL and 1 mL of reservoir solution containing 3.5–4.5 M sodium formate. Large, single crystals grew in 3–7 days at room temperature. Data collection was performed at BioCARS beamline 14-BM-C (λ = 0.90 Å) at the Advanced Photon Source at Argonne National Laboratory. Crystals were flash-frozen directly, with no additional cryoprotectant, in a stream of nitrogen at 100 K. Six data sets were collected, from a total of 4 crystals in the resolution range 5–1.10 Å, in order to maximize data completeness. A seventh data set was collected in the resolution range 28–1.50 Å. This "low-resolution" data set was collected with a beam intensity that was 20% of that used for the high-resolution data in order to accurately measure the high-intensity low-resolution reflections. Data were collected with oscillation widths of either 1° or 0.25° using an ADSC Quantum 4 charge-coupled device (CCD) detector (Area Detector Systems Corporation, Poway, CA). The data were integrated and merged using DENZO and SCALEPACK[21] (Table I).

Structure Refinement

The 1.65-Å structure of FGF-1 [Protein Data Bank (PDB) accession code: 1JQZ] was used as the starting model for refinement (2 molecules per asymmetric unit, and with solvent omitted), and 5% of the reflections were withheld for the calculation of R_{free}.[22] The Crystallography and NMR System (CNS)[23] was used to perform an

initial rigid body refinement, with isotropic temperature factors from the search model, and resulted in values of 31.1% and 31.6% for R_{cryst} and R_{free}, respectively. After several rounds of positional refinement by simulated annealing, followed by water picking, isotropic B-factor refinement, and minor rebuilding using O,[24] R_{cryst} and R_{free} were reduced to 21.6% and 22.9%, respectively. A total of 279 water molecules was added to the structure at this point.

Further refinement was carried out using conjugate gradient least-squares minimization (CGLS) with SHELXL-97.[25] Refinement using isotropic thermal factors resulted in essentially no improvement in the model. Subsequent refinement utilized 6-parameter restrained ADPs instead of individual isotropic thermal factors. This step resulted in a substantial improvement in both R_{cryst} and R_{free} to 16.3% and 19.1%, respectively. The anisotropic treatment of the atoms contributed to improved phases and electron density, which permitted modeling of disorder in the form of several alternative side-chain conformations, as well as more extensive water structure. Water molecules were added using the interactive automated water-picking process available in SHELXL-97. These water molecules were inspected for spherical densities and plausible hydrogen-bonding partners before being included in the model.

In the final stages of refinement, hydrogen atoms were generated by SHELXL-97 according to established stereochemical criteria. These hydrogen atoms were not directly refined but were regenerated before each refinement cycle. Some hydrogen atoms (generally on backbone atoms) were visible in difference maps calculated with the hydrogen atoms removed from the model. The addition of hydrogen atoms resulted in improvements of R_{cryst} and R_{free} to 15.6% and 17.9%, respectively. Final refinement of the data, after modeling all visible alternative conformations, and a total of 407 solvent molecules, resulted in values for R_{cryst} and R_{free} of 14.7% and 17.1%, respectively. This final value of R_{cryst} was calculated from the last round of refinement using all data, including the 5% of reflections previously set aside for cross-validation.

Analysis of Core-Packing Defects

The presence of core-packing defects within the FGF-1 structure was evaluated using the Molecular Surface Package (MSP)[26] and a probe radius of 1.0 Å. Calculations were performed on the refined structure sans discrete hydrogen atoms and utilizing a unified atom model for the van der Waals radii. Cavities identified using a 1.4-Å probe radius can potentially contain disordered solvent, not visible by X-ray diffraction,[27] and thus do not necessarily represent packing "defects." However, the smaller the probe radius, the greater the uncertainty in the calculation due to positional error in the model coordinates. Due to these considerations, a 1.0-Å probe radius was selected for these calculations.

Analysis of ADPs

The equivalent isotropic displacement parameter $\langle U^2 \rangle$ (the mean-square displacement, MSD) is related to the standard isotropic B factor by the relationship

$$B = 8\pi^2 \langle U^2 \rangle \quad (1)$$

The magnitude of U in terms of the refined ADPs[28] is given as

$$U = 1/3(U^{11} + U^{22} + U^{33}) \quad (2)$$

Anisotropy, A, is defined as the ratio of the smallest to the largest eigenvalue of the 3×3 matrix U described by the 6 ADP terms. Both A and U are useful quantities in approximate comparisons of the overall magnitude of the ADP, but they do not provide information regarding the orientation of the anisotropy, and consequently, two atoms with identical values of U and A may have different directions of vibration. The MSD of an atom along a specific vector can be calculated from the ADP terms and by defining the specific vector of interest.[29] If the vector is the interatomic vector between two atoms (bonded or not), and if the two atoms share a rigid body motion, then the difference in the MSD along the interatomic vector will be zero. Correlated ADPs between two atoms A and B, along the interatomic vector, n, quantitate the degree of possible rigid body motion for the atoms as a pair:

$$\Delta_{AB} = < U_A^2 > n \; < U_B^2 > n. \quad (3)$$

A Δ_{AB} value of 0 indicates that the ADPs are identical along the interatomic vector, and this is a necessary condition for the pair of atoms to move as a rigid body. The correlation decreases with the magnitude of the delta value, with negative values indicating the ADPs are negatively correlated. The Δ_{AB} values for all possible pairwise combinations of C^α atoms within a structure can be plotted as a matrix (i.e., a "delta plot"). Alternatively, the absolute values ($|\Delta_{AB}|$) for all pairwise atom combinations can be plotted, and the average computed, to quantify the overall degree of correlated ADPs within a particular region. A delta plot for C^α atoms of FGF-1 was calculated using the program ANISOANL.

TLS Analysis

Although a Δ_{AB} value of 0 is a necessary condition for correlated motions between two atoms, it is an insufficient condition, and by itself does not confirm the presence of rigid body domain motion. Rigid body domains suggested by Δ_{AB} value analysis were evaluated in terms of a TLS model for domain motion.[30,31] TLS analysis was carried out using the program ANISOANL and the TLS-derived ADPs were fitted to the actual ADPs obtained from full anisotropic refinement. Similarity between the derived and actual ADPs can be evaluated using an "R factor" and "goodness-of-fit" (GOF) term:

$$R \text{ factor} = \sqrt{\sum \Delta U^2 / \sum \Delta U_{obs}^2} \quad (4)$$

$$\text{GOF} = \sqrt{\sum \Delta U^2 / (\#\, of\, obs. - \#\, of\, params.)} * 1000. \quad (5)$$

RESULTS

Atomic Resolution Structure of FGF-1

The data quality meets the requirements of "atomic resolution" as defined by Sheldrick and Schneider,[25] in

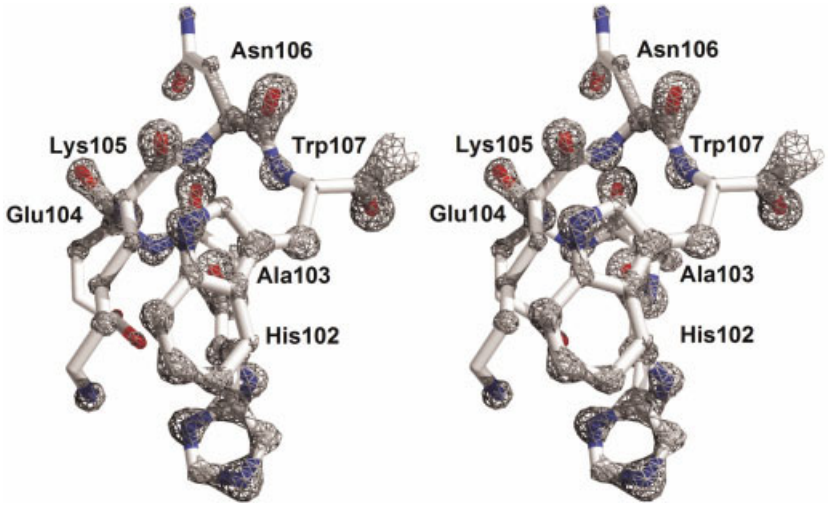

Fig. 2. A relaxed stereo diagram of the atomic model of the refined 1.10-Å structure of human FGF-1 in the region of positions 102–107 (molecule B). Included in the image is the $2F_o$-F_c electron density map contoured at 4.0 σ. This region conveniently illustrates the ability to unambiguously discriminate between the carbon, nitrogen, and oxygen atoms within various side-chains (e.g., His102, Asn106).

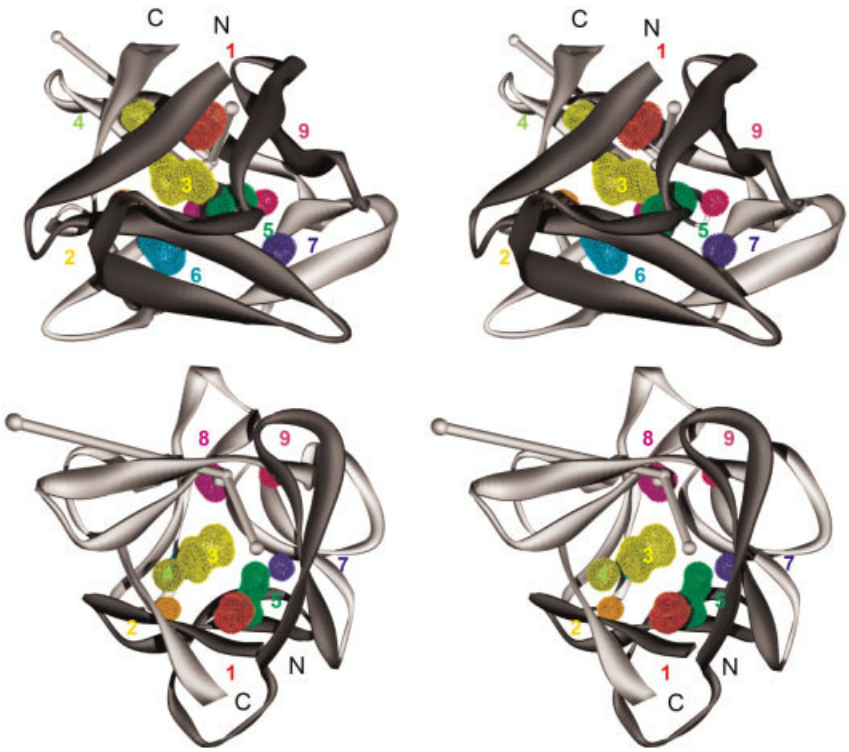

Fig. 3. A relaxed stereo ribbon diagram of FGF-1 as a "side" view (**top**) and looking down the 3-fold axis of structural symmetry (**bottom**) and showing the locations of core-packing defects (cavities) identified using a 1.0-Å probe radius. The cavities are numbered as in Table II, and also color coded, 1–9, from red to violet. The locations of the amino and carboxyl termini are indicated by the letters N and C, respectively. Also included in the figure is the TLS vector associated with the domain involving β-strands 6–12 (shaded as light gray).

that the data extend to at least 1.2-Å resolution, with 50% or more of the theoretically measured reflections in the last resolution bin having $I > 2\ \sigma$. The total number of reflections used by SHELXL-97 for the refinement of the FGF-1 structure was 131,042, and there were 24,420 parameters with 29,744 restraints, yielding a corresponding data-to-parameter ratio of 5.4. The final structure contains 2 molecules in the asymmetric unit and consists of 2294 nonhydrogen protein atoms and 418 solvent atoms (or 152 additional solvent molecules in comparison to the previously reported 1.65-Å structure[20]). As an indication of the increased contrast afforded by the additional terms

of the Fourier series present in the 1.10-Å data, it was possible to discriminate between the nitrogen and carbon atoms on His side-chains, the nitrogen and oxygen atoms of Asn and Gln side-chains, and the stereochemical pucker of Pro side-chains, allowing accurate determination of the orientation of such residues (Fig. 2). The structure factors and refined model of FGF-1 have been deposited in the PDB (accession 1RG8).

Improvements are seen in the Ramachandran plot, with 5 residues moving into the most favored regions (increasing the total to 93.9% from 90.7%). The backbone is well ordered; however, multiple conformations were modeled for the side-chains of Lys57, Lys105, Ser116, and Cys117 of molecule A and Ser116, Cys117, and Leu133 of molecule B. In addition, Pro residues 4 and 11 of molecule A and residues 4, 11, 121, and 136 of molecule B were modeled in two different ring-puckering conformations. The side-chains of Ser17, Lys113, and Lys128 of the A molecule and Lys9, Lys10, Ser17, Gln40, and Arg122 of the B molecule showed significant disorder and could not be modeled in discrete conformations. As with the previously reported 1.65-Å structure, the main-chain atoms of the first two His-tag residues and residues 138–140 of the C-terminus are undefined in the electron density maps.

A final refinement of the structure using blocked full-matrix least-squares mode was performed with no restraints in order to accurately estimate the coordinate error of each atom. The protein was split into 20 blocks of 15 residues, each with a 1-residue overlap. From this method, average coordinate errors were estimated at 0.03 Å for backbone atoms and 0.04 Å for all atoms. The root-mean-square difference between the 1.10-Å model and the previously reported 1.65-Å model main-chain atoms is 0.21 Å, which is significant given the estimated coordinate error of the atomic resolution model.

Extent of Core Packing Defects

A total of 9 cavities was identified within the hydrophobic core of FGF-1 using a 1.0-Å probe radius (Fig. 3). These "microcavities" range in size from 6 to 39 $Å^3$ and have a combined volume of 150 $Å^3$. The largest cavity is centrally located within the structure; however, none of the cavities are detectable using a 1.4-Å probe radius.

Analysis of FGF-1 Disorder Using the ADP Model and TLS Analysis

The equivalent isotropic MSD parameters $\langle U^2 \rangle$ for main-chain atoms of the two independent models of FGF-1 in the asymmetric unit are shown in Figure 4. The values are highly correlated even though the crystal contacts differ between the molecules. Therefore, the $\langle U^2 \rangle$ values reflect an intrinsic property of the molecule and are not a manifestation of the crystal environment or lattice energy. Not surprisingly, the regions with the largest values include residues 5–11 of the N-terminal, as well as turn regions comprising residues 17–20, 27–29, 35–40, 48–51, 59–62, and 113–115. The region encompassing residues 66–109 is well-ordered and displays the lowest equivalent isotropic displacement. Contained within this region is the

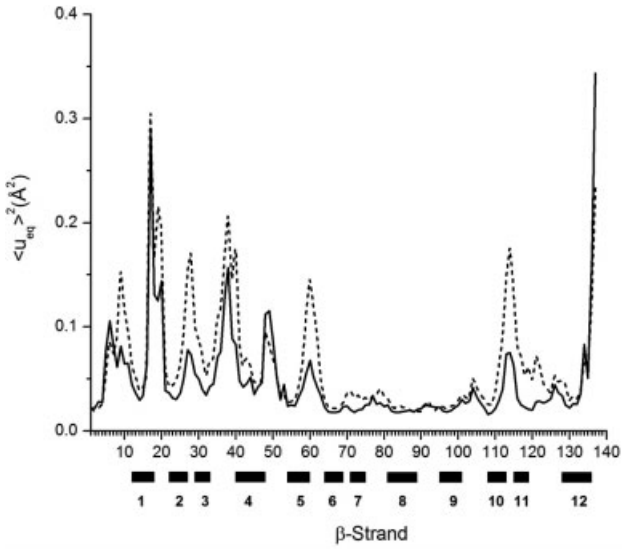

Fig. 4. MSD values for main-chain atoms of FGF-1. There are two independent molecules in the asymmetric unit: molecules A (solid line) and B (dotted line). The locations of the 12-β-sheet secondary structure elements are also shown.

well-ordered β8–β9 turn, which has been identified as a key region for dimerization and receptor binding in different members of the FGF family.[32,33] Generally speaking, highly conserved hydrophobic residues have the lowest individual equivalent isotropic displacement.

The absolute value of the MSD values ($|\Delta_{AB}|$) along the interatomic vector for all pairwise C^{α} combinations in the FGF-1 structure is shown in Figure 5. In order to focus upon the correlations between the β-sheet structural elements, the plot was prepared omitting the turn regions. The $|\Delta_{AB}|$ plot for FGF-1 suggests two general domains within the structure that exhibit correlated anisotropy. The demarcations of these domains are located between β-strands 12 and 1 (the carboxyl and amino termini, respectively) and the other in the region of β-strand 5.

Results of the TLS analysis for various domain definitions comprising 2, 3, 4, 5, 6, or 8 β-strands, the entire structure as a single domain, as well as the β-barrel and lower β-hairpin triplet comprising separate domains, are listed in Supplementary Table SI. The best R factor and GOF values are observed for domain definition #1 for domains defined by two adjacent β-strands, domain definition #3 for domains defined by three adjacent β-strands, domain definition #6 for domains defined by four adjacent β-strands, and domain definition #22 for a domain defined by five adjacent β-strands. For domains defined by six adjacent β-strands, the best R factor and GOF values are observed for domain definition #34; however, there is little difference between the values for any of the six β-strand domain definitions. Since these results consistently identified a demarcation of rigid body motion between β-strands 12 and 1 (see discussion below), the existence of a possible second rigid body boundary was evaluated in the TLS analysis by scanning for a second boundary at all positions between β-strands 2 through 11 (Supplementary Table SII). The results of this analysis indicate that β-strands

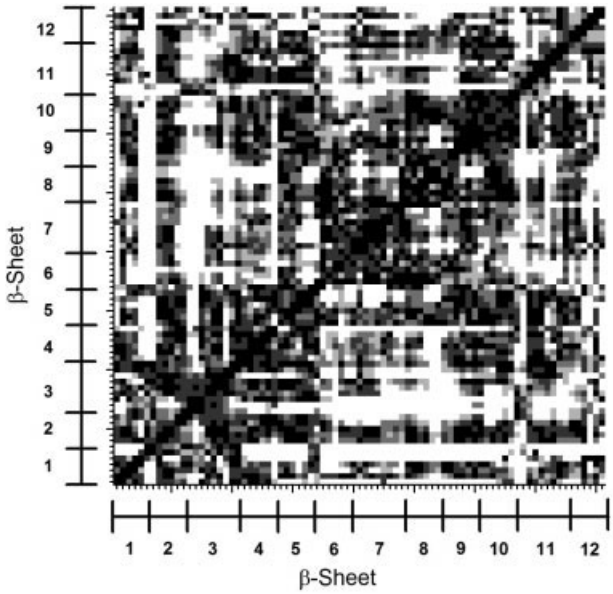

Fig. 5. $|\Delta_{AB}|$ matrix for the refined anisotropic positional displacement along the Cα atom interatomic vectors of FGF-1 (the average values for both molecules within the asymmetric unit is plotted). The values are shaded as follows: black, $0 \leq |\Delta_{AB}| < 0.25\ \sigma$; dark gray, $0.25\ \sigma \leq |\Delta_{AB}| < 0.50\ \sigma$; gray, $0.50\ \sigma \leq |\Delta_{AB}| < 0.75\ \sigma$; light gray, $0.75\ \sigma \leq |\Delta_{AB}| < 1.00\ \sigma$; white, $1.00\ \sigma \leq |\Delta_{AB}|$, where the σ values are calculated from the set of β-sheets in the structures.

6–12 exhibit correlated rigid body motion; however, β-strands 1–5 as a group appear to be somewhat flexible and do not form a correlated rigid body. The vectors of the TLS libration tensors associated with the β-strands 6–12 domain are shown in Figure 3, and the displacement around the longest libration axis is 3.2°[2].

DISCUSSION

The FGF-1 structure reported here is the highest resolution structure determined to date for any member of the β-trefoil superfold. This structure provides data on the absolute configuration of Asn, Gln, and His residues; pucker of Pro residues; more accurate modeling of residues with multiple conformations; more extensive solvent structure; lower coordinate error; and information regarding atomic anisotropic displacement. FGF-1 exhibits relatively low thermal stability,[34,35] and a previous 1.65-Å X-ray structure of FGF-1 identified the presence of four cavities within the core region using a 1.0-Å probe radius.[20] However, it was noted also that the volume calculations of such cavities are sensitive to small positional errors in the coordinates of the model. The refinement of an atomic resolution structure of FGF-1 provides an opportunity to re-evaluate such internal cavities with greater certainty. The results of the cavity analysis using the present 1.10-Å atomic resolution structure indicate the presence of nine cavities within the core region of FGF-1 (Table II). These nine cavities comprise the original four seen in the 1.65-Å structure, as well as five additional small cavities. Cavity #3 (the largest) is located at the center of the structure and is a characteristic feature of all β-trefoil structures (although the volume varies substantially for different members). Of the remaining eight cavities, six are associated with the first trefoil subdomain (Table II). Of these six cavities associated with trefoil 1, four are at the interface between trefoil 1 and the rest of the structure, and two are located internally to trefoil 1. This analysis therefore shows that the region of the structure associated with the largest number of internal cavities is the first trefoil domain (β-strands 1–4).

The presence of internal cavities does not necessarily mean that the core-packing volume is less than optimal. Only if such defects can be filled without the introduction of strain can it be concluded that there is a more optimal set of core-packing residues. Serrano and coworkers[36] have shown, through a series of core-packing mutants, that "overstuffing" the core can contribute a favorable free-energy term due to the hydrophobic effect, but also an unfavorable term due to the introduction of strain. In some cases, the two energetic terms can largely cancel each other out, resulting in a mutant with stability essentially indistinguishable from the wild-type. However, such mutants can be readily distinguished from the wild-type by examining their folding and unfolding kinetic constants. The additional buried hydrophobic area increases the folding rate constant, whereas the introduction of strain increases the unfolding rate constant. The net result is that the mutant will exhibit a "chevron plot" that is shifted in the vertical axis in comparison to the wild-type protein. Therefore, a characteristic property of a less-than-optimal core-packing volume is one where greater buried area can be introduced, but without an associated introduction of strain (i.e., without an increase in the rate of unfolding). We have reported an extensive study on core-packing mutants of FGF-1, including folding and unfolding kinetic data, and involving a combination mutant that results in a net increase in the core volume of four methyl groups.[16] Although this mutant was designed from a consideration of increasing the primary structure symmetry, it was noted that the additional methyl groups should reside within several of the identified central cavities. Various biophysical analyses showed that this mutant is associated with no distortions of the tertiary structure, a reduction in the packing defects detectable using a 1.0-Å probe radius, and no increase in the rate of unfolding (i.e., no increase in structural strain). Thus, we conclude that the core of FGF-1 has a combined volume that is less than optimal (at least with regard to biophysical considerations) and the identified cavities represent packing defects.

Protein dynamics is a key aspect in understanding protein function; however, it can be one of the most challenging properties of proteins to characterize. It has long been recognized that proteins with more than one structurally compact subdomain can exhibit independent motions of these domains.[37] Such motions have been inferred from comparisons of different crystal forms of the same protein, ADP and TLS analysis of atomic-resolution X-ray structures, covariation of NMR-derived backbone dynamical parameters, and molecular dynamics simulations.[38–41] Core-packing defects have been proposed as the

TABLE II. Volume and Location of Solvent-Excluded Cavities Within FGF-1 Detectable Using a 1.0-Å probe Radius.

Cavity	Volume (Å^3)	Center (X,Y,Z)			Neighbor Positions	β-strands	Trefoil(s)
1	17	16.5	−38.4	36.1	13, 14, 44	1, 4	1
2	6	17.9	−31.8	40.2	14, 22, 111, 132	1, 2, 10, 12	1, 3
3	39	12.8	−34.9	41.5	(Center of structure)		
4	8	11.8	−34.0	35.8	14, 95, 97, 132	1, 9, 12	1, 3
5	26	17.9	−38.9	42.5	23, 25, 44	2, 4	1
6	20	16.0	−32.6	45.8	23, 73, 110, 117	2, 7, 10, 11	1, 2, 3
7	7	17.9	−41.6	46.6	25, 31, 65, 74	2, 3, 6, 7	1, 2
8	22	8.5	−35.7	47.2	67, 85, 99, 109	6, 8, 9, 10	2, 3
9	6	8.4	−42.0	47.5	83, 85	8	

The neighboring residues to the cavities, as well as their associated β-strand and trefoil, are also listed.

basis of the observed rigid body motions for the structurally compact subdomains of the NG domain of *Thermus aquaticus* Ffh protein, identified using atomic-resolution X-ray data with ADP refinement and TLS analysis.[41]

In a recent study of TIM-barrel enzymes in the histidine biosynthesis pathway, Wilmanns and coworkers[42] have provided evidence that this fundamental superfold evolved via a gene duplication and fusion event. In this case, the two structural subunits are not visible as independent, compact structural elements, but integrate seamlessly to form the TIM barrel superfold. An analysis of *apo*- and *holo*-forms of 2,5-diketo-D-gluconate reductase, an aldo-keto reductase sharing the TIM barrel superfold, has identified a hinge-bending motion between 2 halves of the protein in response to binding of β-nicotinamide adenine dinucleotide phosphate (NADPH) cofactor.[43] This hinge-bending motion demarcates the aforementioned duplicated region. Thus, it is apparent that domain motions can occur for symmetrically related regions within a symmetric superfold, even though such domains are not readily identifiable as structurally compact separate domains.

The plot of the MSD values for FGF-1 suggests that the carboxyl-terminal region of the protein, consisting of approximately β-strands 6–12, exhibits generally lower positional displacements in comparison with the amino-terminal region (Fig. 4). Analysis of the anisotropic positional displacement correlations along the C^{α} interatomic vectors suggests a potential rigid body demarcation between the termini and within the region of β-strand 5 (Fig. 5). As previously mentioned, $\Delta_{AB} = 0$ is a necessary but insufficient condition for defining rigid body motions. Confirmation of rigid body motions can be provided by TLS analysis and evaluation of the resulting *R* factor and GOF values.[31,41,44] For the various domains evaluated, involving groupings of 2, 3, 4, or 5 β-strands, a consistent pattern is present in the *R* factor and GOF values of the TLS analysis: The greatest correlations are observed for definitions where β-strand 12 is not paired with β-strand 1 (Supplementary Table SI). Thus, the TLS results are consistent and confirm that the termini in FGF-1 demarcate a boundary of rigid body motion. The question arises as to whether there is another discernible demarcation of rigid body motion within the molecule (possibly involving β-strand 5, as suggested by the delta plot) or whether such motion is distributed around the entire structure. Scanning along the structure with a second demarcation in the TLS analysis shows that β-strands 6–12 comprise a rigid body domain (Supplementary Table SII). However, these results also suggest that individual β-strands 1–5 do not exhibit significant correlated anisotropy with each other; thus, this region appears as a relatively flexible separate subdomain. The TLS results obtained for both independent molecules of FGF-1 were consistent. We note also that while the β-trefoil structure can be described as a 6-stranded β-barrel, capped at one end by 3 β-hairpins, the TLS analysis for FGF-1 does not support any rigid body motion for such subdomain definitions.

The identified rigid body subdomain in FGF-1, involving β-strands 6–12, includes 3 of the 4 β-strands of trefoil 2 and all of trefoil 3. Thus, the domain motion does not neatly follow the putative gene duplication regions. However, the β-hairpin structures within FGF-1 (i.e., the pairs of β-strands stabilized by the largest number of inter-strand hydrogen bonds) involve β-strands 2–3, 4–5, 6–7, 8–9, and 10–11 (Fig. 1). These pairings highlight the circularly permutable, domain-swapped characteristics of the β-trefoil superfold and suggest why the carboxyl-terminal rigid body domain does not include β-strand 5.

The cellular receptors for the FGF family of proteins characteristically contain 3 extracellular immunoglobulin (Ig)-like domains, 2 of which interact with bound FGF. However, signaling requires that 2 such complexes dimerize so as to activate the intracellular tyrosine kinase portion of the receptor. Formation of the dimer complex subsequently requires the presence of heparin, a polysulfonated polysaccharide present in the extracellular matrix or as heparan on cell-surface glycoproteins, forming a signaling-competent FGF/FGF receptor (FGFR)/heparin ternary complex. Although two different FGF structures (FGF-1 and FGF-2) for the ternary complex have been reported,[45,46] the details of the individual FGF/FGFR interactions and FGF/heparin interactions share some similarities. There are 14 close contacts (i.e., within 3.4 Å) between the receptor Ig-like domains and FGF in the ternary complex (PDB accession: 1FQ9). Of these, 10 are located within the first trefoil subunit, 1 is in trefoil 2, and 3 are located within trefoil 3. In contrast, there are 8 close contacts between the heparin and FGF in the ternary complex, 1 of which is located in trefoil 1, none in trefoil 2, and 7 are located in trefoil 3. The receptor-binding and

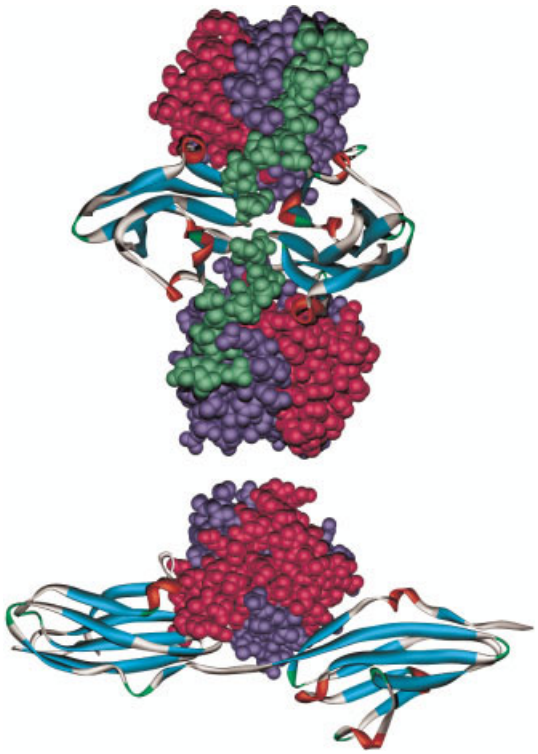

Fig. 6. FGF/FGFR/heparin ternary complex.[45] FGF is shown in space-filling representation with the rigid body domain comprising β-strands 1–5 indicated in red, and the domain comprising β-strands 6–12 in blue. The heparin polysaccharide component of the complex is shown in green space-filling representation. The two Ig-like ligand-binding domains of the FGFR are rendered in a ribbon representation. The **upper panel** shows the ternary complex, while the **lower panel** shows a single FGF/FGFR complex to highlight the interaction of β-strands 1–5, with the two Ig-like domains in a single receptor molecule. The lower panel is rotated approximately 90° about the vertical axis in relationship to the upper panel.

heparin-binding functionalities within FGF-1 are therefore segregated within different regions of the structure. In this regard, the carboxyl-terminal rigid body domain (β-strands 6–12) contains essentially the entire set of heparin-binding contacts, while the amino-terminal domain (β-strands 1–5) contains the majority of receptor-binding contacts (Fig. 6). Flexibility between these functionally independent regions may play an essential role in the productive formation of a stable, signaling-competent, ternary complex of FGF, FGFR, and heparin. Rigidity of the heparin-binding domain may also be an essential structural feature that contributes to the binding affinity for specific heparin sulfate glycosaminoglycans.[47] Furthermore, FGF-1 is unique among the FGFs in that it is the only known member that can bind to all known FGF receptors. A recent structure of the FGF10/FGF receptor 2b structure[48] shows that this complex requires an unusual rotation of one of the receptor Ig-like domains, forming a previously undescribed conformation for an FGF with its receptor. Thus, the ability of FGF-1 to interact with all known FGF receptors would likely require a certain structural plasticity. An important question is why a protein that exhibits a clear structural symmetry should exhibit asymmetric domain motions. The fact is, despite the tertiary structure symmetry (i.e., 3-fold) present in FGF-1, this symmetry is essentially absent at the level of the primary structure, and this contributes to the asymmetric features of the core packing defects. We continue to study this question.

ACKNOWLEDGMENTS

We thank Dr. Mohammad Yousef for technical assistance with SHELXL-97 and the staff at BioCARS for technical assistance with data collection. We are also grateful to Dr. Ethan Merritt for helpful discussions and to Dr. Martyn Winn for helpful discussions concerning TLS analysis.

REFERENCES

1. Zhu X, Komiya H, Chirino A, Faham S, Fox GM, Arakawa T, Hsu BT, Rees DC. Three-dimensional structures of acidic and basic fibroblast growth factors. Science 1991;251:90–93.
2. Sweet RM, Wright HT, Janin J, Chothia CH, Blow DM. Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 angstrom resolution. Biochemistry 1974; 13:4212–4228.
3. Tahirov TH, Lu TH, Liaw YC, Chen YL, Lin JY. Crystal structure of abrin-a at 2.14Å. J Mol Biol 1995;250:354–367.
4. Rutenber E, Katzin BJ, Ernst S, Collins EJ, Mlsna D, Ready MP, Robertus JD. Crystallographic refinement of ricin to 2.5Å. Proteins 1991;10:240–250.
5. Lacy DB, Tepp W, Cohen AC, DasGupta BR, Stevens RC. Crystal structure of botulinum neurotoxin type A and implications for toxicity. Nat Struct Biol 1998;5:898–902.
6. Emsley P, Fotinou C, Black I, Fairweather NF, Charles IG, Watts C, Hewitt E, Isaacs NW. The structures of the H(C) fragment of tetanus toxin with carbohydrate subunit complexes provide insight into ganglioside binding. J Biol Chem 2000;275:8889–8894.
7. Priestle JP, Schar H-P, Grutter MG. Crystallographic refinement of interleukin 1β at 2.0 angstrom resolution. Proc Natl Acad Sci USA 1989;86:9667–9671.
8. Habazettl J, Gondol D, Wiltscheck R, Otlewski J, Schleicher M, Holak TA. Structure of hisactophilin is similar to interleukin-1β and fibroblast growth factor. Nature 1992;359:855–858.
9. Kaneko S, Kuno A, Fujimoto Z, Shimizu D, Machida S, Sato Y, Yura K, Go M, Mizuno H, Taira K, Kusakabe I, Hayashi K. An investigation of the nature and function of module 10 in a family F/10 xylanase FXYN of *Streptomyces olivaceoviridis* E-86 by module shuffling with the *Cex* of *Cellulomonas fimi* and by site-directed mutagenesis. FEBS Lett 1999;460:61–66.
10. McLachlan AD. Three-fold structural pattern in the soybean trypsin inhibitor (Kunitz). J Mol Biol 1979;133:557–563.
11. Murzin AG, Lesk AM, Chothia C. β-Trefoil fold: patterns of structure and sequence in the Kunitz inhibitors interleukins-1β and 1α and fibroblast growth factors. J Mol Biol 1992;223:531–543.
12. Mukhopadhyay D. The molecular evolutionary history of a winged bean α-chymotrypsin inhibitor and modeling of its mutations through structural analysis. J Mol Evol 2000;50:214–223.
13. Ponting CP, Russell RB. Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all beta-trefoil proteins. J Mol Biol 2000;302:1041–1047.
14. Zhang J, Cousens LS, Barr PJ, Sprang SR. Three-dimensional structure of human basic fibroblast growth factor, a structural homolog of interleukin 1B. Proc Natl Acad Sci USA 1991;88:3446–3450.
15. Blaber M, DiSalvo J, Thomas KA. X-ray crystal structure of human acidic fibroblast growth factor. Biochemistry 1996;35:2086–2094.
16. Brych SR, Kim J, Logan TM, Blaber M. Accommodation of a highly symmetric core within a symmetric protein superfold. Protein Sci 2003;12:2704–2718.
17. Harper JW, Lobb RR. Reductive methylation of lysine residues in acidic fibroblast growth factor: effect on mitogenic acitivity and heparin affinity. Biochemistry 1988;27:671–678.

18. Patrie KM, Botelho MJ, Franklin K, Chiu IM. Site-directed mutagenesis and molecular modeling identify a crucial amino acid in specifying the heparin affinity of FGF-1. Biochemistry 1999;38: 9264–272.
19. Seddon AP, Aviezer D, Li L-Y, Bohlen P, Yayon A. Engineering of fibroblast growth factor: alteration of receptor binding specificity. Biochemistry 1995;34:731–736.
20. Brych SR, Blaber SI, Logan TM, Blaber M. Structure and stability effects of mutations designed to increase the primary sequence symmetry within the core region of a β-trefoil. Protein Sci 2001;10:2587–2599.
21. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
22. Brunger AT. Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 1992;355:472–475.
23. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54: 905–921.
24. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 1991;A47: 110–119.
25. Sheldrick GM, Schneider TR. SHELXL-97. Methods Enzymol 1997;277:319–343.
26. Connolly ML. The molecular surface package. J Mol Graphics 1993;11:139–141.
27. Yu B, Blaber M, Gronenborn AM, Clore GM, Caspar DLD. Disordered water within a hydrophobic protein cavity visualized by X-ray crystallography. Proc Natl Acad Sci USA 1999;96:103–108.
28. Willis BTM, Pryor AW. Thermal vibrations in crystallography. London: Cambridge University Press; 1975. 296 p.
29. Schneider TR. What can we learn from anisotropic temperature factors? In: Proceedings of the CCP4 study weekend. Dodson E, Moore M, Ralph A, Bailey S, editors. Daresbury, UK: SERC Daresbury Laboratory; 1996. p 133–144.
30. Schomaker V, Trueblood KN. On the rigid-body motion of molecules in crystals. Acta Crystallogr B 1968;24:63–76.
31. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr D Biol Crystallogr 2001;57:122–133.
32. Hecht HJ, Adar R, Hofmann B, Bogin O, Weich H, Yayon A. Structure of fibroblast growth factor 9 shows a symmetric dimer with unique receptor- and heparin-binding interfaces. Acta Crystallogr D Biol Crystallogr 2001;57:378–384.
33. Pellegrini L, Burke DF, von Delft F, Mulloy B, Blundell TL. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature 2000;407:1029–1034.
34. Copeland RA, Ji H, Halfpenny AJ, Williams RW, Thompson KC, Herber WK, Thomas KA, Bruner MW, Ryan JA, Marquis-Omer D, Sangal G, Sitrin RD, Yamazaki S, Middaugh CR. The structure of human acidic fibroblast growth factor and its interaction with heparin. Arch Biochem Biophys 1991;289:53–61.
35. Blaber M, Adamek DH, Popovic A, Blaber SI. Biophysical and structural analysis of human acidic fibroblast growth factor. In: Marshak DR, editor. Techniques in protein chemistry. Vol. VIII. San Diego: Academic Press; 1997. p 745–753.
36. Ventura S, Vega MC, Lacroix E, Angrand I, Spagnolo L, Serrano L. Conformational strain in the hydrophobic core and its implications for protein folding and design. Nat Struct Biol 2002;9:485–493.
37. Ptitsyn OB. Inter-domain mobility in proteins and its probable functional role. FEBS Lett 1978;93:1–4.
38. Faber HR, Matthews BW. A mutant T4 lysozyme displays five different crystal conformations. Nature 1990;348:263–266.
39. Gibrat JF, Go N. Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. Proteins 1990;8:258–279.
40. Mayer KL, Earley MR, Gupta S, Pichumani K, Regan L, Stone MJ. Covariation of backbone motion throughout a small protein domain. Nat Struct Biol 2003;10:962–965.
41. Ramirez UD, Minasov G, Focia PJ, Stroud RM, Walter P, Kuhn P, Freymann DM. Structural basis for mobility in the 1.1Å crystal structure of the NG domain of *Thermus aquaticus* Ffh. J Mol Biol 2002;320:783–799.
42. Lang D, Thoma R, Henn-Sax M, Sterner R, Wilmanns M. Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. Science 2000;289:1546–1550.
43. Sanli G, Dudley JI, Blaber M. Structural biology of the aldo-keto reductase family of enzymes: catalysis and cofactor binding. Cell Biochem Biophys 2003;38:79–101.
44. Howlin B, Butler SA, Moss DS, Harris GW, Driessen HPC. TLSANL: TLS parameter-analysis program for segmented anisotropic refinement of macromolecular structures. J Appl Crystallogr 1993;26:622–624.
45. Plotnikov AN, Schlessinger J, Hubbard SR, Mohammadi M. Structural basis for FGF receptor dimerization and activation. Cell 1999;98:641–650.
46. DiGabriele AD, Lax I, Chen DI, Svahn CM, Jaye M, Schlessinger J, Hendrickson WA. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. Nature 1998;393:812–817.
47. Raman R, Venkataraman G, Ernst S, Sasisekharan V, Sasisekharan R. Structural specificity of heparin binding in the fibroblast growth factor family of proteins. Proc Natl Acad Sci USA 2003;100: 2357–2362.
48. Yeh BK, Igarashi M, Eliseenkova AV, Plotnikov AN, Sher I, Ron D, Aaronson SA, Mohammadi M. Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. Proc Natl Acad Sci USA 2003;100:2266–2271.