

**Crystallographic studies on the Fibroblast Growth
Factors Protein Family
And
Cys 112 Asp azurin from *Pseudomonas aeruginosa***

**Thesis by
Salem Faham**

**In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy**

*California Institute of Technology
Pasadena, California*

1998

(Submitted August 8, 1997)

Dedicated to my family

Acknowledgment

I would like to thank my advisor Professor Doug Rees for his support and guidance throughout my graduate years. Also, I would like to thank Xiaotian Zhu, Barbara Hsu, Michael Chan, and John Peters for their friendship and help. And, finally, I would like to thank my family for their love.

Abstract

The biological activities of fibroblast growth factors (FGFs) are profoundly influenced by interactions with heparin-like molecules. Crystal structures of basic FGF (bFGF) complexed with heparin-derived tetra- and hexa-saccharides have been determined at resolutions of 1.9Å and 2.2Å, respectively. Both heparin fragments bind in similar fashions to a region of the bFGF surface containing residues Asn-28, Arg-121, Lys-126, and Gln-135; the heparin hexamer additionally interacts with an adjacent site formed by Lys-27, Asn-102, and Lys-136. No significant conformational change in bFGF occurs upon heparin binding, suggesting that heparin serves to facilitate the juxtaposition of components of the FGF signal transduction pathway. Comparison of the binding sites for various anions to bFGF indicates that while many anions bind to this general region of bFGF, a wide variation in specific binding interactions can be achieved. Combined with sequence variability in this region between different members of the FGF family, it is likely that a diverse set of binding interactions between FGFs and heparin-like molecules occur, which may be reflected in the differing responses and requirements for these sulfated polysaccharides exhibited by the FGF family.

Keratinocyte growth factor (KGF) is a member of the fibroblast growth factors (FGF) protein family, which stimulate the proliferation and differentiation of a variety of cell types. The crystal structure of KGF has been solved at 2.4 Å resolution by molecular replacement, using basic fibroblast growth factor (bFGF) as a starting model. The model refined to a final R factor of 21.0 and R_{free} of 32.%. The protein has the same fold as that of the other FGF family members. New features are found in this structure, including a disulfide bond and different conformations for some of the loops. The differences in the

structure are likely to be responsible for the differences in heparin and receptor binding requirements between the different FGF proteins, thereby identifying regions of possible functional importance.

Replacement of the cysteine at position 112 of *Pseudomonas aeruginosa* azurin with an aspartic acid residue results in a mutant (Cys112Asp) protein that retains a strong copper binding site. Cu^{II} (Cys112Asp) azurin can be reduced by excess $[\text{Ru}^{\text{II}}(\text{NH}_3)_6]^{2+}$, resulting in a Cu^{I} protein with an electronic absorption spectrum very similar to that of wild-type Cu^{I} azurin. Cys112Asp azurin exhibits reversible interprotein electron-transfer reactivity with *P. aeruginosa* cytochrome *c*₅₅₁ ($\mu = 0.1$ M sodium phosphate (pH 7.0)); $E^\circ(\text{Cu}^{\text{III/I}}) = 180$ mV vs. NHE); this redox activity indicates that electrons can still enter and exit the protein through the partially solvent-exposed imidazole ring of His117. The structure of Cu^{II} (Cys112Asp) azurin at 2.4 Å resolution shows that the active-site copper is five coordinate: the pseudo-square base of the distorted square-pyramidal structure is defined by the imidazole N^δ atoms of His46 and His117 and the oxygen atoms of an asymmetrically-bound bidentate carboxylate group of Asp112; the apical position is occupied by the oxygen atom of the backbone carbonyl group of Gly45. The Cu^{II} -Asp112 interaction is distinguished by an approximately 1-Å displacement of the metal center from the plane defined by the Asp112 carboxylate group.

Contents

Acknowledgement		iii
Abstract		iv
Chapter 1	Introduction to the Fibroblast Growth Factors (FGF) Protein Family	1
Chapter 2	Crystal Structures of bFGF complexed with Heparin tetra- and hexa- saccharides	14
	Supplemental Material	38
Chapter 3	Crystal Structure of Keratinocyte Growth Factor	46
Appendix	Role of the Active-Site Cysteine of <i>Pseudomonas aeruginosa</i> Azurin. Crystal Structure Analysis of the Cu^{II}(Cys112Asp) Protein	59

List of Figures

Chapter 1.

Fig. 1.1.	Chemical structure of the repeating disaccharide of heparin.	3
Fig. 1.2.	Sequence alignment of the twelve human FGF proteins. Homologous residues are shaded in gray.	8
Fig. 1.3.	Schematic representation of the FGF receptor. IG, IG2, and IG3 represent the extracellular immunoglobulin-like domains; TM, transmembrane region; TK1 and TK2, tyrosine kinase domains.	10
Fig. 1.4.	Ribbons diagram of aFGF showing the β barrel fold adopted by members of the FGF protein family.	10

Chapter 2.

Fig. 2.1.	Structures of the heparin tetra- and hexa- saccharides used for the crystallographic studies. The numbering system used here to identify the different sugar units is indicated, with the nonreducing ring assigned the number 1.	15
Fig. 2.2.	Electron density map for the heparin hexamer, contoured at twice the overall rms value of the map.	17
Fig. 2.3.	Superposition of the tetrasaccharide and hexasaccharide heparin fragments binding to bFGF.	20
Fig. 2.4.	The heparin binding site, illustrating the specific interactions between bFGF and the hexasaccharide fragment of heparin.	21
Fig. 2.5.	The 1C_4 chair and 2S_0 skew boat ring conformations of L-iduronic acid. The nomenclature describes the conformation of the six membered hexopyranose ring (C or S for chair and skew boat, respectively), while the super- and sub-scripted numbers refer to the two ring atoms most displaced from the plane defined by the other four ring atoms	25

Fig. 2.6.	Superposition of the aFGF-sucrose octasulfate and bFGF-heparin hexasaccharide structures.	28
------------------	---	----

Chapter 3.

Fig. 3.1.	Harker section $W=1/2$ of the native Patterson map of the KGF crystals showing the packing peak at (0.36,0.22,0.5). The Patterson function was calculated using data between 20.0 to 6.0 Å resolution.	48
Fig. 3.2.	Ribbons diagram of the three dimensional structure of KGF.	50
Fig. 3.3.	Superposition of KGF with aFGF. KGF is in light gray.	51
Fig. 3.4.	Ribbons diagram of KGF showing in black the loops that have the largest structural differences from aFGF and bFGF.	54
Fig. 3.5.	Stereodiagram of KGF with regions I, and II in black, and showing the position of heparin hexamer after superpositioning with the bFGF-heparin structure.	56
Fig. 3.6.	Stereodiagram of KGF showing in black loops that are likely to have structural variations due to size differences among members of the FGF family. The position of heparin after the superposition with the bFGF-heparin structure is shown.	57

Appendix

Fig. A.1.	Stereodiagram of the active site of C112D azurin showing a Fourier difference map ($F_o - F_c$) at 3.0 overall rms value, where residue 112 was deleted from the model.	70
Fig. A.2.	Ribbons diagram of the Cu^{II} (Cys112Asp) azurin dimer that constitutes an asymmetric unit.	71
Fig. A.3.	Stereodiagram of the amino termini of native (black) and Cys112Asp azurin (gray), showing the different conformations adopted by the first two residues, the hydrogen bonding between the amino nitrogen and the hydroxyl group of Ser25, and the coordination geometry of the exogenous copper.	72
Fig. A.4.	Stereodiagram of the coordination structure of the active-site	73

copper of Cu^{II}(Cys112Asp) azurin (gray) superimposed on the active site of the native azurin structure (black). See Table A.1 for a list of interatomic distances.

List of Tables

Chapter 2.

Table 2.1.	Polar (ion pair and hydrogen bond) contacts between bFGF residues and tetrasaccharide and hexasaccharide fragments of heparin. Interactions between atoms separated by less than 3.6Å are indicated.	33
Table 2.2.	Values of the ϕ , ψ torsion angles about the glycosidic bonds in the hexasaccharide fragment of heparin complexed to bFGF. ϕ and ψ are defined as the torsion angles about the H1-C1-O4'-C4' and the C1-O4'-C4'-H4', respectively, that connect successive sugar rings by 1,4 linkages.	34
Table 2.3.	Sequence alignment of members of the FGF family in the heparin binding region. The GCG package was used to generate the alignment. Heparin residues contacting side chain atoms of the indicated residues are listed, with the abbreviations G = glucosamine and I = iduronic acid. High affinity refers to contacts involving GlcN-2 or Idu-3, while low affinity refers to contacts involving Idu-5 or GlcN-6. The accession numbers for the bFGF, aFGF, INT2, HST, FGF5, FGF6, KGF, AIGF and FGF9 sequences are A26642, A24820, S04742, Tvhuhs, Tvhuf5, S04204, A36301, A46245, and A48137, respectively, of the PIR database.	35

Appendix

Table A.1.	Active-Site Interatomic Distances for <i>P. aeruginosa</i> Cu ^{II} (Cys112Asp) and WT Azurins.	69
-------------------	---	----

Chapter 1

Introduction to the Fibroblast Growth Factors (FGF) Protein Family

Introduction

Members of the fibroblast growth factor (FGF) family stimulate the growth and differentiation of a wide variety of cell types through receptor mediated signal transduction pathways (reviewed in (1,2,3,4,5)). An unusual aspect of this process is the participation of sulfated polysaccharides in the initial binding events between FGFs and receptors. The high affinity of FGFs for sulfated polysaccharides such as heparin or heparan sulfate is a distinguishing feature of the FGF family, and has been extensively characterized for the first two members of the FGF family, acidic FGF (aFGF or FGF-1) and basic FGF (bFGF or FGF-2). There is a general consensus that, at a minimum, heparin-like molecules enhance the activity of FGFs either through stabilization of the protein structure or through binding-induced conformational changes (6). There is also evidence that heparin-like molecules can directly participate in FGF - receptor interactions, either by binding to both components, or by binding multiple FGF molecules to facilitate receptor dimerization (7,8,9,10). Aside from their role in mediating the FGF - receptor interaction, heparan sulfate proteoglycans in the extracellular matrix may also function to stabilize FGFs, or to help sequester and localize FGFs in the vicinity of the cell surface.

These functional properties reflect sequence specific structural features of heparin and other sulfated proteoglycans that permit productive binding interactions with FGFs. The heparin polymer is generated from a basic disaccharide repeat unit consisting of L-iduronic acid and D-glucosamine joined by α 1-4 linkages (Fig. 1.1). A typical heparin disaccharide contains a total of three sulfate groups; one attached to the 2-hydroxyl group of iduronic acid and two linked to the 2-amino and 6-hydroxyl groups of glucosamine. Heparan sulfate found in cell surface proteoglycans has a similar structure, although it is

typically less extensively sulfated than heparin. X-ray fiber diffraction studies indicate that heparin adopts a helical structure with the disaccharide as the fundamental structural repeat unit. Successive disaccharides within heparin are related by a twofold screw operation, with a rotation angle of $\sim 180^\circ$ and a translation of $\sim 8.5 \text{ \AA}$ along the helix axis (11). Although the average structure of heparin is well defined, local structural alterations can exist within the heparin polymer that may be important for interactions with other biomolecules. These structural alterations arise from two sources: (i) variability in the sulfation pattern that can modulate the structure and affinity of heparin for other molecules and (ii) variability in the ring conformation of the L-iduronic acid. Three different low energy ring conformations are potentially available for iduronic acid residues (12), and the distribution between these conformations is sensitive to the size of the heparin molecule, the identity and degree of sulfation of neighboring sugar units, and the presence of counterions (13). Consequently, binding of heparin to other molecules, including FGFs, may be accompanied by conformational changes within the polysaccharide chain.

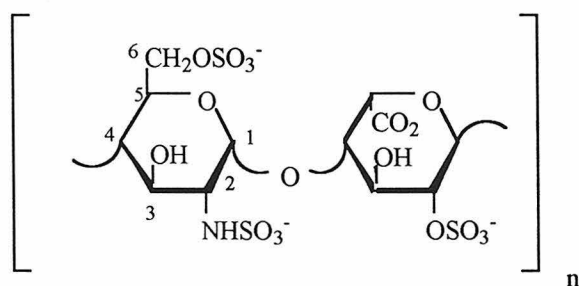


Fig 1.1. Chemical structure of the repeating disaccharide of heparin.

A variety of experimental studies have implicated specific regions of both FGFs and heparin as participating in the binding interface between these two molecules. These observations may be summarized:

Binding site for heparin on FGFs:

The participation of residues in the 120-140 region of bFGF (the numbering system of bFGF described in (14) is used throughout the thesis) in heparin binding has been established by chemical modification (15), proteolysis (16), and peptide competition experiments (17); the latter studies are also consistent with a role for residues 25-29. Site directed mutagenesis studies have demonstrated that substitutions in residues K27, N28, R82, K120, R121, T122, Q124, K126, K130, Q135, and K136 interfere with heparin binding (18). Crystallographic studies of FGFs have identified several of these residues (K120, R121, K126 and K130) as involved in the binding of anionic groups (19,20,21,22,23). The crystal structure of aFGF-sucrose octasulfate (SOS) complex has provided a detailed structure of the interaction between a sulfated disaccharide and residues corresponding to K120, T122, Q124, K126, K130, and Q135 in bFGF (24). More recently, a structure of bFGF complexed with a non-sulfated heparin trisaccharide analog was reported (25). Two molecules were found to bind FGF, one at the known heparin binding site interacting with N28, K120, R121, K126, and K136, while the other made contacts with residues R98, E97, and R61. Spectroscopic analysis of the interaction between heparin and bFGF suggests that heparin binding induces a conformational change in the protein structure that might be relevant to the activation of FGFs by heparin (26,27), although only small localized changes in the FGF structure have been observed crystallographically upon binding of FGF to other sulfated ligands.

Heparin-protein interactions are often characterized as driven primarily by ionic forces (28). When the energetics of the complex formation between bFGF and heparin were analyzed, however, only 30% of the binding energy was attributable to ionic interactions (18). This study further indicated that three sodium ions are displaced upon binding and that eight H-bonds may form between heparin and bFGF. The dissociation constant for the binding of heparin to bFGF in 0.1M NaCl was measured as $\sim 0.5 \mu\text{M}$.

Binding interactions with heparin:

From the influence of partial desulfation of heparin on binding affinity, the N-sulfate group of glucosamine and the 2 O-sulfate group of iduronic acid appear to be essential for bFGF binding, while the 6 O-sulfate group of glucosamine apparently is not required (29). A size dependence on the binding of heparin fragments to FGF has been observed; while tetra- or penta-saccharide heparin fragments can bind to FGF (29,30,31), the biological activation of the FGF requires a longer fragment. Full activation is achieved with heparin fragments composed of 14 saccharide units, although some activity is observed with fragments of 12 saccharide units (31,32,33). Binding studies indicate that one bFGF molecule can bind per octasaccharide sequence (34), while aFGF binds at a density of one molecule per 4-5 saccharide units (35).

The FGF family consists of twelve members (FGF 1 through 12) (36), with homologies ranging from 30% to 55% (Fig. 1.2). Despite the overall similarities between members of the FGF protein family, the details of their mode of action is different.

Currently, four different FGF receptors (FGFR) are identified (FGFR 1 to 4). Different splice forms are also generated by the FGFR genes. The FGFR is composed of an intracellular domain with tyrosine kinase activity, putative transmembrane helix, and

an extracellular domain consisting of two to three immunoglobulin domains (Fig. 1.3.). It is believed that activation of the FGFR occurs by ligand induced receptor dimerization. The FGF receptors bind members of the FGF family with a varying degree of specificity. KGF interacts solely with one receptor, the KGF receptor (KGFR), which is identical to FGFR-2 except for a stretch in the carboxy end of the Ig3 domain. Both KGFR and FGFR-2 are generated from the same gene, and differences are a result of the splicing process. A pairwise study on the interactions between nine FGF members (FGF1-9) and all the FGFR's has revealed that aFGF is the only FGF member capable of activating all FGFR splice variants (37). It has also confirmed that KGFR binds both aFGF and KGF well, but not bFGF, and that FGFR-2 binds both bFGF and aFGF well, and not KGF. It has been deduced that KGF is recognized by the Ig3 domain of its receptor (KGFR). Keratinocyte growth factor is an unusual member of the FGF protein family, in that it only acts upon cells of epithelial origin. This indicates that its receptor is expressed only in epithelial cells. However, KGF is solely produced by a cells of mesenchymal origin (38), suggesting that KGF is a mediator of epithelial-mesenchymal communications, a feature shared only with FGF3. Due to this special feature, it has been suggested that KGF might lead to potential wound healing agents (39). The crystal structures of aFGF and bFGF revealed a β barrel structure made of 12 β strands (Fig. 1.4.). KGF with about 34% sequence identity to both aFGF and bFGF is expected to have the same fold.

To establish a structural framework for probing the interactions between heparin and FGF, we have initiated a crystallographic analysis of this system. Since heparin itself is too heterogeneous to be compatible with crystallization, a series of oligosaccharide

fragments of heparin have been prepared by heparinase digestion and cocrystallized with bFGF. In this thesis, the structures of the binding of tetra- and hexa-saccharide fragments of heparin to FGF are reported (chapter 2) (40). While generally consistent with many of the previous studies on the binding of heparin to FGF, these structures provide a precise definition of the heparin geometry and binding mode to FGF, and emphasize the structural variability that can be accommodated in the heparin structure and through members of the FGF family. Crystallographic analysis of the KGF structure was undertaken as well (chapter 3), to better understand the structural factors responsible for the functional variability of the various FGF members.

```

FGF1> -----
FGF2> -----
FGF3> -----MGLIWLLLLSLL
FGF4> -----MSGPGTAAVALLPAVLLALAPWAGRGGAAAPTAPNGTLEAELE
FGF5> -----MSLSFLLLLFFSHLI SAWAHGEKRLAPKGQPGPAATDRN
FGF6> MALGQKLFITMSRGAGRLQGTLWALFLGIIVGMVVPSPAGTRANNTLLD
FGF7> -----MHKWILTWLPTLLYRSCFHII CLVGTISLACN
FGF8> -----MGSPRSASCLLLHLLVLCLQAQEGPGRGPAL
FGF9> -----APLGEVGNVYFGVQDAVPFGNVPVL
FGF10> -----
FGF11> -----
FGF12> -----

```

```

FGF1> -----AEGEITTTFT-----ALTEKFN---LPPN-----YKPPPL
FGF2> -----AAGSITTLP-----ALPEDGGSGAFPPH-----FADPPT
FGF3> EPGWPAAGPGARLRR-----DAGGRGGVYEHGGA-----PPTPL
FGF4> RR-WESVALSLARLPVAAQPKEAAVQSGAGDYIL-----IPLLE
FGF5> PIGSSSRQSSSSAMSSSSASSSPAASLGSQSGEQSSFQWSPSGRTGS
FGF6> SRGWGTLSRSRAGL---AG--EIAGVNWESGYVE-----IQQRI
FGF7> DMTPEQATNVNCSSP-----ERHTRSVDYEEG-----DIQVRI
FGF8> GRELASFRAGREPQG--VSQQHVREQSLVTDQSRR-----LIETIQ
FGF9> PVDSPVLSDLHGQSE--AGGLPRGPAVTDLDHKEI-----LEEIQ
FGF10> -----MASKEP-----QLKGIVT-----
FGF11> -----
FGF12> -----

```

```

FGF1> EYNSG-HFRLEPQTDENR RSDQHIOQSAESAE-EYINSTEI
FGF2> EYNSG-HFRHEDGRDVRKSDPHIKQQAEEER-EYINSTEI
FGF3> EYATK--YHIOHSGFNGLNS-AGHLETAEE-EYINSTEI
FGF4> EYNVIGIHHIALHGRGGAHAT-RDGLLESPER-EYINSTEI
FGF5> EYNVIGIHHIYFGKNHGHAN-MLVLEFAEQ-EYINSTEI
FGF6> EYNVIGIHHIQLHGRGSGHHEEN-PSGLLESTER-EYINSTEI
FGF7> EYNTQ--YRDKRGNKQMKNNNLEERTAA-EYINSTEI
FGF8> EYNTSG-KVQVLANKENMAEGDPAAKIEETDTFSRRRPAET
FGF9> EYNT--YHLEFNNTQQRKHHSRGNLEFISAE-EYINSTEI
FGF10> EYQQ--FLQHEPDTDEKKNSDYFNIPGRMAQQA
FGF11> EYQQ--FLQHEPDTDEKKNSDYFNIPGRMAQQA
FGF12> EYQQ--FLQHEPDTDEKKNSDYFNIPGRMAQQA

```

```

FGF1> EYQVAMDTDELK--QTPNNTLLELEHHEI-KKH-----
FGF2> EYQVAMDTDELK--KCVTDEFFERESNNTNRK-----
FGF3> EYQVAMDTDELK--EYHAEVERHGLGTAARRTVSSTPG
FGF4> EYQVAMSSKGLY--PFTDETFEGLLPNMAEYK-----
FGF5> EYQVAMSSKGLY--AKTDDKTRERFQNSHMAAH-----
FGF6> EYQVAMSSKGLY--PSQLLEKRETLNMAED-----
FGF7> EYQVAMSSKGLY--KKECNENNGLELEHHEI-AKKT-----
FGF8> EYQVAMSSKGLY--KNGKGDVTEIILENNTALQNAK-----
FGF9> EYQVAMSSKGLY--SEKLQEQVEHQFEETWNTSNT-----
FGF10> EYQVAMSSKGLY--SDVETPEKKEKSFEEYVLSST-----
FGF11> EYQVAMSSKGLY--SPHETAEKREKCFEEYVLSST-----
FGF12> EYQVAMSSKGLY--SDVETPEKKEKSFEEYVLSST-----

```

```

FGF1> -----AEINFEVGGKNSCTGPP--THYGVAILLPLP--SSD---
FGF2_> -----YTSFPPAKKTQYLLS--TGPGVAILLPLMS-AKS----
FGF3_> ARQPSAEILLQYSLGKRPFGF--TSTTQSLPLPLV--DHRDHE
FGF4_> -----PQMFAISNKTGNGN--VSPTMVTHTLPP--I-----
FGF5_> --TEKTSEETVALNREKANGCSPV--PQHISTHTLPPF-KQSEQPE
FGF6_> -----QTYLAISLYRVVGG--VSPIMTVHTLPLI-----
FGF7_> -----HNGEMEVAFCKEIPVPEK--TTEETATLPLMA--IT----
FGF8_> -----EETVTFTEKCRPPTES--TQHQTEVETKELPRGHHTTE
FGF9_> --HVDTEPRYPALENDTPEET--TCHHNFHTLPLP--DPDKVP
FGF10_> --QQESGRANWELGLNFEQIMNGN--VE--TKPSSHTVPP--LEVCMYR
FGF11_> --QRRSGRANWELGLDDEQVMNGN--V--TKAAAHTLPL--L-----
FGF12_> --QQESGRANWELGLNFEQIMNGN--V--TKPSSHTVPP-----

```

```

FGF1> -----
FGF2_> -----
FGF3_> MVRQLQSGLP RP PGKVQPRRRRQKQSPDNLEPSHVQASRLGSQLEASAH
FGF4_> -----
FGF5_> LSFTVTVP EKKNPSP I KSKIPLSAPRKNTNSVKYRLKFRFG-----
FGF6_> -----
FGF7_> -----
FGF8_> QSLRFEFLNYP PFTRSLRGSQRTWAPEPR-----
FGF9_> ELYKDILSQS-----
FGF10_> EPSLHEIGENKGVQ GKFWTP-----
FGF11_> -----
FGF12_> -----

```

```

FGF1> -
FGF2_> -
FGF3_> -
FGF4_> -
FGF5_> -
FGF6_> -
FGF7_> -
FGF8_> -
FGF9_> -
FGF10_> P
FGF11_> -
FGF12_> -

```

Fig. 1.2. Sequence alignment of the 12 human FGF proteins. Conservatively

substituted residues are shaded in gray.

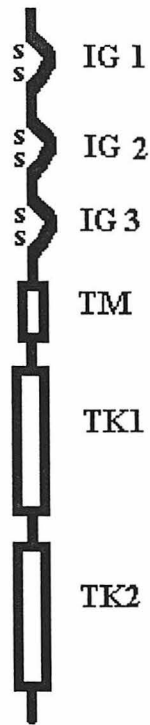
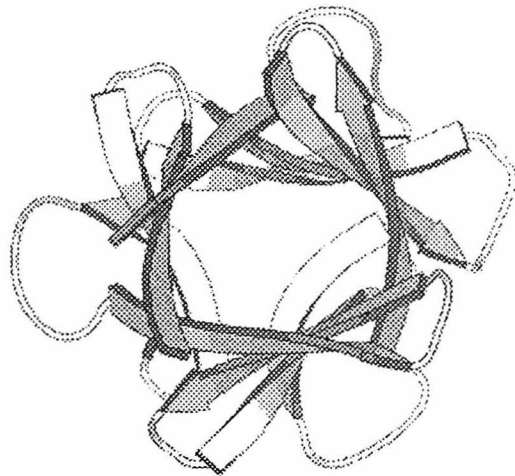


Fig. 1.3. Schematic representation of the FGF receptor. IG 1, IG 2, and IG 3 represent the extracellular immunoglobulin-like domains; TM, transmembrane region; TK1 and TK2, tyrosine kinase domains.

Fig. 1.4. Ribbons diagram of aFGF showing the β barrel fold adopted by members of the FGF protein family (23).



References

1. D. Gospodarowicz, N. Ferrara, L. Schweigerer, G. Neufeld, *Endocrinol. Rev.* **8**, 95-114 (1987).
2. W. H. Burgess, T. Maciag, *Annu. Rev. Biochem.* **58**, 575-606 (1989).
3. D. Benharroch, D. Birnbaum, *Isr. J. Med. Sci.* **26**, 212-219 (1990).
4. C. Basilico, D. Moscatelli, *Adv. Cancer Res.* , 115-165 (1992).
5. J. Folkman, Y. Shing, *J. Biol. Chem.* **267**, 10931-10935 (1992).
6. I. J. Mason, *Cell* **78**, 547-552 (1994).
7. D. M. Ornitz, et al., *Mol. Cell. Biol.* **12**, 240-247 (1992).
8. T. Spivak-Kroizman, et al., *Cell* **79**, 1015-1024 (1994).
9. M. Kan, et al., *Science* **259**, 1918-1921 (1993).
10. S. Guimond, M. Maccarana, B. B. Olwin, U. Lindahl, A. C. Rapraeger, *J. Biol. Chem.* **268**, 23906-23914 (1993).
11. I. Nieduszynski, in *Heparin* D. A. Lane, U. Lindahl, Eds. (CRC Press, Boca Raton, Florida, 1989) pp. 51-63.
12. P. N. Sanderson, T. N. Huckerby, I. A. Nieduszynski, *Biochem. J.* **243**, 175-181 (1987).
13. D. R. Ferro, A. Provasoli, M. Ragazzi, *Carbohydr. Res.* **195**, 157-167 (1990).
14. G. M. Fox, et al., *J. Biol. Chem.* **263**, 18452-18458 (1988).
15. J. W. Harper, R. R. Lobb, *Biochemistry* **27**, 671-678 (1988).
16. R. R. Lobb, *Biochemistry* **27**, 2572-2578 (1988).
17. A. Baird, D. Schubert, N. Ling, R. Guillemin, *Proc. Natl. Acad. Sci. USA* **85**, 2324-2328 (1988).
18. L. D. Thompson, M. W. Pantoliano, B. A. Springer, *Biochemistry* **33**, 3831-3840 (1994).

19. J. Zhang, L. S. Cousens, P. J. Barr, S. R. Sprang, *Proc. Natl. Acad. Sci. USA* **88**, 3446-3450 (1991).
20. H. Ago, Y. Kitagawa, A. Fujishima, Y. Matsuura, Y. Katsube, *J. Biochem.* **110**, 360-363 (1991).
21. A. E. Eriksson, L. S. Cousens, L. H. Weaver, B. W. Matthews, *Proc. Natl. Acad. Sci. USA* **88**, 3441-3445 (1991).
22. A. E. Eriksson, L. S. Cousens, B. W. Matthews, *Protein Science* **2**, 1274-1284 (1993).
23. X. Zhu, et al., *Science* **251**, 90-93 (1991).
24. X. Zhu, B. T. Hsu, D. C. Rees, *Structure* **1**, 27-34 (1993).
25. D. M. Ornitz, et al., *Science* **268**, 432-436. (1995).
26. S. J. Prestrelski, G. M. Fox, T. Arakawa, *Arch. Biochem. Biophys.* **293**, 314-319 (1992).
27. R. A. Copeland, et al., *Arch. Biochem. Biophys.* **289**, 53-61 (1991).
28. A. D. Lander, *Chem. & Biol.* **1**, 73-78 (1994).
29. M. Maccarana, B. Casu, U. Lindahl, *J. Biol. Chem.* **268**, 23898-23905 (1993).
30. D. B. Volkin, et al., *Arch. Biochem. Biophys.* **300**, 30-41 (1993).
31. A. Walker, J. E. Turnbull, J. T. Gallagher, *J. Biol. Chem.* **269**, 931-935 (1994).
32. J. E. Turnbull, D. G. Fernig, Y. Ke, M. C. Wilkinson, J. T. Gallagher, *J. Biol. Chem.* **267**, 10337-10341 (1992).
33. D. Aviezer, et al., *J. Biol. Chem.* **269**, 114-121 (1994).
34. T. Arakawa, J. Wen, J. S. Philo, *Arch. Biochem. Biophys.* **308**, 267-273 (1994).
35. H. Mach, et al., *Biochem.* **32**, 5480-5489 (1993).
36. A. S. Verdier, M. G. Mattei, H. Lovec, H. Harung, M. Goldfarb, D. Birnbaum, *Genomics* **40**, 151-154 (1997).
37. D. M. Ornitz, J. S. Xu, J. S. Colvin, D. G. Mcewen, C. A. Macarthur, F. Coulier, G. X. Gao, M. Goldfarb, *J. Biol. Chem.* **271**, 15292-15296 (1996).

38. J. S. Rubin, D. P. Bottaro, M. Chedid, T. Miki, D. Ron, H.-G. Cheon, W. G. Taylor, E. Fortney, H. Sakata, P. W. Finch, W. J. LaRochelle, *Cell Bio. Int.* **19**, 399-411 (1995).
39. C. Sotozono, T. Inatomi, M. Nakamura, S. Kinoshita, *Inv. Ophth. & Vis Sci.* **36**, 1524-1529 (1995).
40. S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt, D.C. Rees, *Science* **271**, 1116-1120 (1996).

Chapter 2

Crystal Structures of bFGF complexed with Heparin tetra- and hexa- saccharides

(modified from Faham et al. , Science 271, 1116 (1996))

Experimental Procedures

Tetrasaccharide and hexasaccharide fragments of heparin were prepared in the laboratory of R. Linhardt, Univ. of Iowa, by partial digestion of bovine lung heparin with heparin lyase I and by preparative SAX-HPLC as described previously (1). As a product of this preparative procedure, the heparin fragments contain an unsaturated uronic acid group at the non-reducing end and a mixture of α/β anomers at the reducing end of the oligosaccharide. Consequently, the tetra- and hexa-saccharide fragments contain one and two intact units of the heparin disaccharides, respectively, in their interior (Fig 2.1).

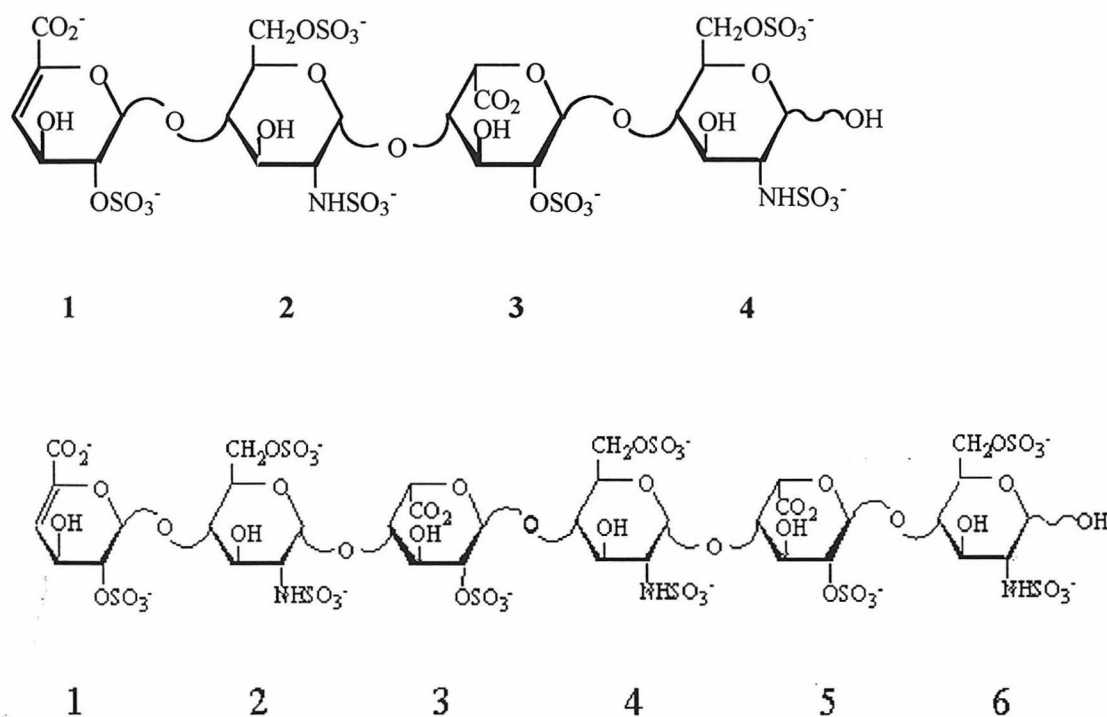


Fig. 2.1. Structures of the heparin tetra- and hexa- saccharides used for the crystallographic studies. The numbering system used here to identify the different sugar units is indicated, with the nonreducing ring assigned the number 1.

Cocrystals of bFGF-heparin tetrasaccharide were grown using the vapor diffusion method, using 1.0 μ l of 20mg/ml heparin tetrasaccharide solution added to 2.5 μ l of a 10mg/ml bFGF solution and 2.5 μ l of reservoir solution. This corresponds to about 10:1 molar ratio of tetrasaccharide to bFGF. The bFGF used in the sample is recombinant human bFGF with cysteines at residues 70 and 88 replaced by serines (2). The mixture was then allowed to equilibrate against the reservoir, which contained 25% PEG8K, and 0.1M *N*-[2-acetamido]-2-iminodiacetic acid (ADA) 7.0. Crystals grew over a few days at room temperature reaching a final size of approximately 0.2mmx0.2mmx0.2mm. The space group was determined to be P2₁2₁2₁ with cell dimensions of a=32.0Å, b=41.8Å, c=85.9Å. Diffraction data were collected at room temperature using a Rigaku RAXIS-IIc imaging plate, and integrated using MSC programs and scaled and merged with CCP4 (3). The data set used for the structure determination contained 9,084 unique reflections to 1.9Å resolution (95.6% complete), merged from a total of 23,651 observations with a merging R of 0.071.

Likewise, cocrystals of bFGF-heparin hexasaccharide were grown using the vapor diffusion method, using 2.0 μ l of 10mg/ml heparin hexasaccharide solution added to 2.5 μ l of a 10mg/ml bFGF solution and 2.5 μ l of reservoir solution. This corresponds to about 6:1 molar ratio of hexasaccharide to bFGF. The mixture was then allowed to equilibrate against the reservoir, which contained 34% PEG8K, 30mM MgCl₂, and 0.1M ADA 7.0. Crystals grew over a few days at room temperature reaching a final size of approximately 0.1mmx0.2mmx0.2mm. The space group was determined to be P2₁2₁2₁ with cell dimensions of a=33.1Å, b=41.9Å, c=86.1Å. Diffraction data was collected at room temperature using a Siemens area detector at room temperature. Data was processed and scaled using the program XDS (4). The data set used for the structure determination contained 6,123 unique reflections to 2.2Å resolution (94.7% complete), merged from a total of 19,016 observations with a merging R of 0.073.

This crystal form had been previously observed for unliganded bFGF crystals crystallized from ammonium sulfate (5), so that molecular replacement methods were not required to solve these structures. Difference Fourier maps calculated using the corresponding unliganded bFGF structure immediately revealed the position of the heparin fragments. Rigid body refinement and subsequent positional and temperature factors refinement were performed using the X-PLOR package (6). New parameter, and topology files were prepared for the refinement of the heparin fragments (supplemental material). The heparin fragments were built into the difference Fourier map using the TOM/FRODO graphics program (7). An $F_o - F_c$ difference electron density map calculated omitting the hexasaccharide from the structural model is illustrated in Fig. 2.2.

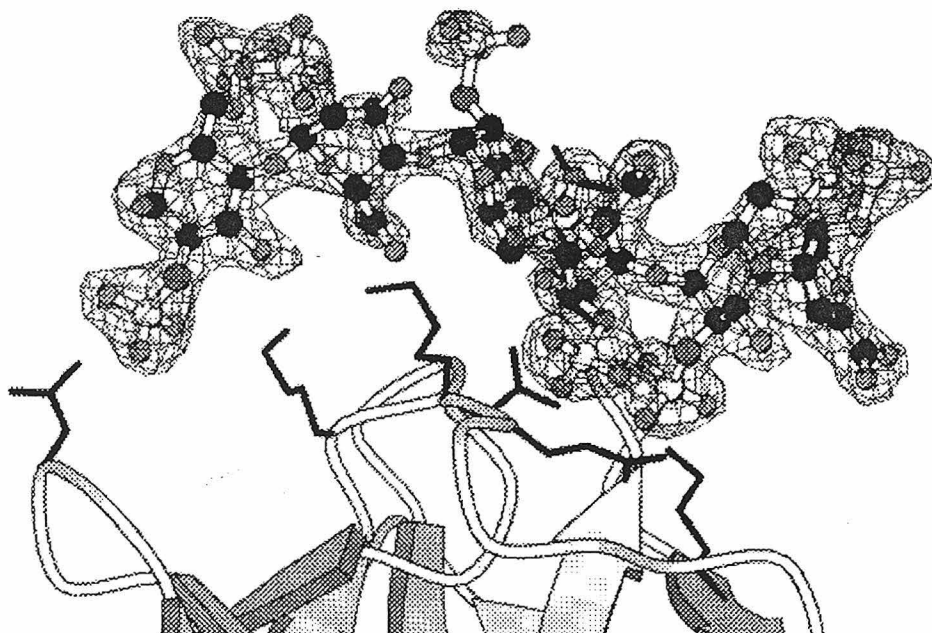


Fig. 2.2. Electron density map for the heparin hexamer, contoured at twice the overall rms value of the map.

In the tetrasaccharide case, the terminal sugar unit at the nonreducing end had a somewhat weaker electron density, indicating its somewhat disordered nature in the crystal. Iterative cycles of refinement and model building were performed for both structures. The bFGF-heparin tetrasaccharide structure refined to $R=0.21$ for data between 5.0 and 1.9Å resolution, with $F > \sigma(F)$, and rms deviations for bond distances and angles from ideal values of 0.015Å and 2.45° respectively. The bFGF-heparin hexasaccharide structure refined to $R=0.21$ for data between 6.0 to 2.2Å resolution, with $F > \sigma(F)$, and rms deviations for bond distances and angles from ideal values of 0.016Å and 2.55° respectively. A total of 1068 atoms are found in the current bFGF-tetramer model assembled from 989 protein atoms, 70 heparin atoms, and 9 water atoms, and a total of 1108 atoms are found in the current bFGF-hexamer structure assembled from 989 protein atoms, 105 heparin atoms, and 14 water atoms.

Superimpositions of the different structures were based on the backbone C_{α} positions and were done using a program based on the algorithm of Kabsch (8). Helical parameters were determined as described by Cox (9) from the rotation and translation parameters obtained (8) for the superposition of the ring atoms of Idu 1-GlcN 2-Idu 3 onto Idu 3 - GlcN 4 - Idu 5. Similar parameters were also derived for superposition of GlcN 2 onto GlcN 4; Idu 1-GlcN 2 onto Idu 3-GlcN 4; and GlcN 2-Idu 3 onto GlcN 4-Idu 5. Structural figures were prepared with the program MOLSCRIPT (10).

Results

Binding of Heparin Fragments to bFGF

Although both complexes were prepared for crystallization using solutions with an excess of the heparin fragment (with ratios of heparin fragment to bFGF molar ratios of 6:1 and 10:1 for the hexa- and tetra-saccharide, respectively), the crystalline complex contained a 1:1 molar ratio of fragment to bFGF in both cases. Using the numbering scheme for the sugar rings defined in Fig. 2.1, rings 1 to 4 of the heparin fragments bind in a very similar fashion to bFGF. Both fragments interact with residues N28, R121, K126, and Q135, while the longer hexasaccharide also contacts residues K27, N102, and K136 (Fig. 2.3). Since the shorter tetrasaccharide and the longer hexasaccharide overlap in the former region, this must represent a high affinity site for the binding of FGF to heparin fragments based on the disaccharide sequence repeat of heparin. As described in chapter 1, this region of bFGF has been implicated in the binding of heparin by a variety of studies.

Many of the specific interactions between bFGF and the heparin hexamer are mediated by negatively charged groups on the heparin fragment (Table 2.1). In particular, the 2-N-sulfate groups of glucosamine residues 2 and 6, the 2 O-sulfate group of iduronic acid 3, and the carboxylate group of iduronic acid residue 5 are involved in salt bridge and hydrogen bonding interactions with polar groups on bFGF. The N-sulfate group on ring 2 forms a salt bridge with Lys126, and hydrogen bonds to the side chain of Asn28, and to the amide NH of residue 121. The N-sulfate group on ring 6 forms a salt bridge with the side chain of Lys27, and hydrogen bonds to the side chain of Asn102. The 2 O-sulfate on ring 3 also forms a salt bridge with the side chain of Lys126, as well as hydrogen bonds to the side chains of Gln135, Asn28, and backbone NH groups of

residues 136 and 137. The carboxylate group on ring 5 forms a salt bridge with Lys136. Additional hydrogen bonding interactions are observed between the 3-OH group of L-iduronic acid ring #3 and the side chains of Arg121 and Asn28,

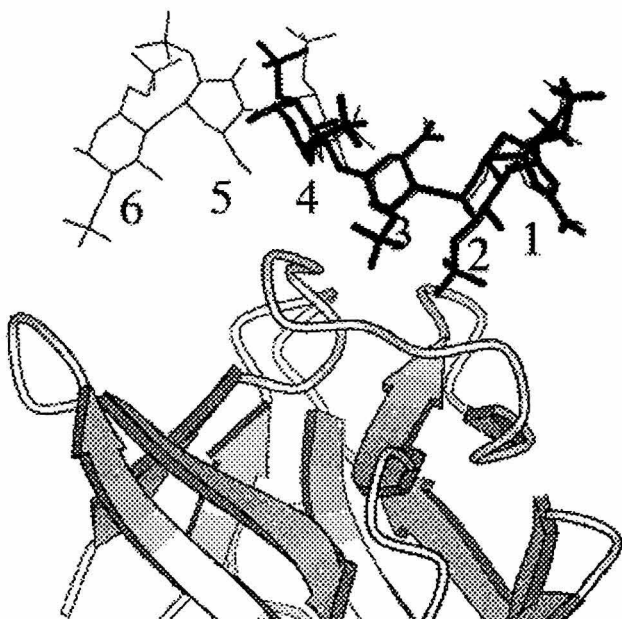


Fig. 2.3. Superposition of the tetrasaccharide and hexasaccharide heparin fragments binding to bFGF.

while the 3-OH group on ring 6 hydrogen bonds to the side chain of Lys27. In total, 9 hydrogen bonds and 6 salt bridges are observed between the heparin hexamer and bFGF (Table 2.2; Fig. 2.4).

The surface area lost by bFGF residues buried upon binding of the tetrasaccharide and the hexasaccharide to bFGF may be calculated to be $\sim 180\text{\AA}^2$ and $\sim 290\text{\AA}^2$,

respectively. For comparison, $\sim 250\text{\AA}^2$ of aFGF surface was buried upon the binding of sucrose octasulfate (11). Only a few residues contribute the bulk of the surface area buried in these complexes; in the tetrasaccharide complex, over half of the buried surface

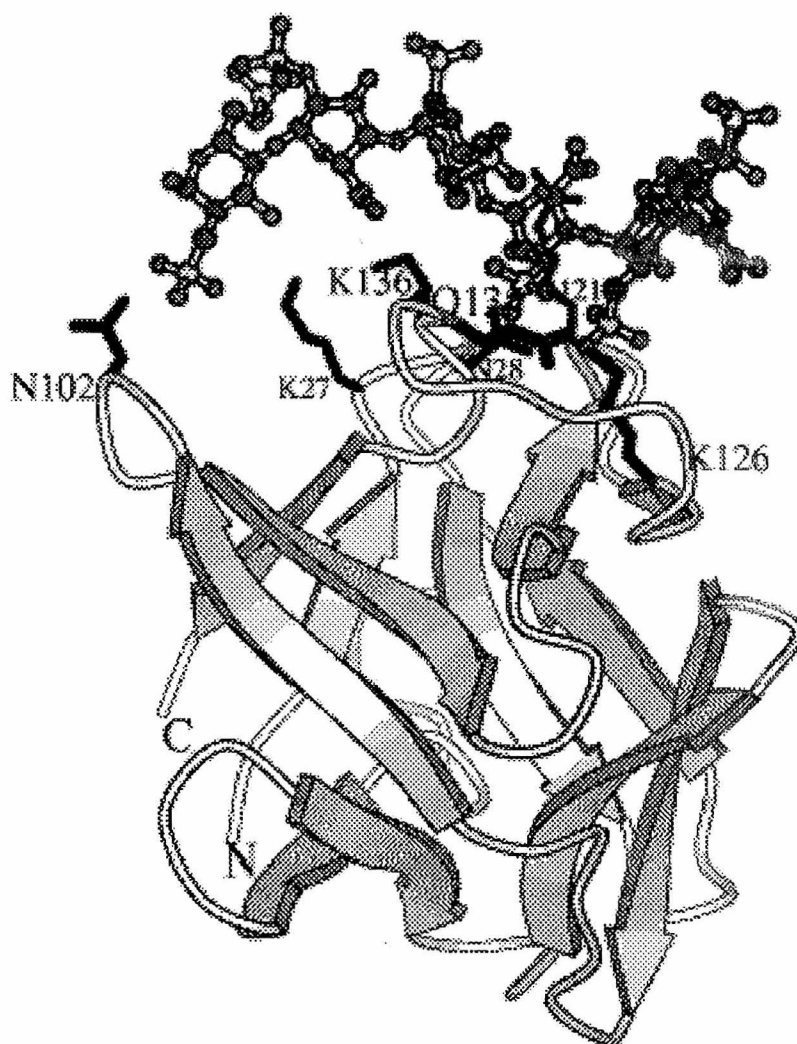


Fig. 2.4. The heparin binding site, illustrating the specific interactions between bFGF and the hexasaccharide fragment of heparin.

area is contributed by three residues, Arg-121, Lys-126 and Lys-136, while Lys-27 and Asn-102 each provide an additional $> 30\text{\AA}^2$ buried surface area in the hexasaccharide complex. The surface area buried in the complex with the hexasaccharide is likely representative of the surface area that would be covered in the binding of bFGF to an extended heparin polymer, since computational extension of the hexamer appears to provide most of the interactions that a linear heparin helix would have with a bFGF monomer. The binding of these heparin fragments to bFGF is not associated with any significant conformational change in the polypeptide backbone; the root mean square (rms) deviation based on the C_α positions between the heparin hexamer-bFGF complex and the unbound bFGF is $0.34\text{\AA}$. Similar observations have been reported for the binding of sucrose octasulfate (11) and nonsulfated trisaccharides to FGFs, as well (12). Some changes in sidechain conformations upon heparin binding are observed including the ordering of the Arg-121 sidechain which only has poor electron density in the unliganded structure.

In addition to the primary contacts described above, the heparin hexamer is also involved in some crystal packing interactions. Those include salt bridges between the carboxylate group of ring 1 and the side chain of Arg108, the 6-O sulfate group on ring 2 with Lys 47 and Arg 23, and the 6-O sulfate on ring 6 with Arg34 and His36. Hydrogen bonds are also generated between the 2-O sulfate group on ring 1 with Asn 72, and the 2-O sulfate on ring 6 with backbone NH of residues 45 and 46. Ring 4 does not make any contacts with bFGF, and in the tetrasaccharide case it has weak electron density and high temperature factors indicating its somewhat disordered nature in that crystal form. Ring 1 is only making contacts with symmetry related molecules. There is no evidence for heparin-fragment induced dimerization of bFGF in these cases, since the areas involved in lattice contacts is relatively small, and since a similar crystal form may be prepared in the absence of heparin fragments.

Structure of the heparin fragments

The two heparin fragments have essentially the same structure, and make the same contacts with bFGF when possible (Fig. 2.3). The helical parameters of the heparin helix have been estimated from the hexamer structure; the average observed helical parameters of a 174° rotation and an $8.6\text{\AA}$ translation per disaccharide are in good agreement with the x-ray fiber diffraction values of 180° and $8.5\text{\AA}$, respectively, for the heparin polymer. The observed torsion angles (Table 2.3) of the heparin hexamer fall within the calculated allowed region for the heparin polymer, and are close to those previously derived from NMR data (13). The only apparent difference in the polysaccharide conformation between the two fragments is a slight shift in the position of GlcN-4. Since GlcN-4 is the reducing end of the tetrasaccharide case and the electron density for this residue somewhat diffuse, the apparent flexibility of this region of the tetrasaccharide might contribute to the observed variability in the glycosidic bond geometry.

A stereochemically significant aspect of the heparin structure concerns the conformation of the six membered hexopyranose ring of the iduronic acid residues. In general, six membered organic rings, exemplified by cyclohexane, will tend to adopt a conformation that minimizes both angle strain within the ring and steric clash between substituents. If possible, for example, cyclohexane will tend to adopt a "chair" type conformation with ring substituents in equatorial positions. Due to the spatial arrangement of substituents about the iduronic acid ring, however, a single lowest energy structure for this residue cannot be uniquely achieved. Instead, iduronic acid residues primarily exist in an equilibrium between two types of ring conformations; a chair form designated 1C_4 that has little bond angle strain in the ring but unfavorable contacts between substituents, and a skew boat conformation termed 2S_0 that reduces the

unfavorable contact between substituents at the expense of increased ring strain (Fig. 2.5). The hexasaccharide fragment of heparin used in this study contains two iduronic acid rings, at positions 3 and 5 (the sugar ring at position 1 is actually an unsaturated Idu residue that contains a Δ 4,5 double bond within the six membered ring). Idu-3 of the hexamer is observed to be in a 1C_4 chair conformation, while Idu-5 adopts the 2S_0 skew boat conformation. It is likely that interactions with the protein are responsible for influencing the observed L-iduronic acid ring conformations. This possibility is supported by the observation that Idu-3 adopts the same conformation in both the tetra- and hexa-saccharide cases despite different fragment lengths, while Idu-5 of the hexasaccharide has a different conformation from ring 3 of the tetrasaccharide, even though these residues both neighbor a terminal glucosamine. Although the global effects of the different conformations of iduronic acid on the structure of the heparin polymer are relatively small, the local effects on binding interactions to FGFs can be significant, since the positions of the ring substituents will differ between the two structures. Consequently, multiple conformations of iduronic acid can be adopted in the binding of heparin to FGF, and that the conformational flexibility of this residue may play a role in the recognition and binding processes of heparin to its various biological ligands.

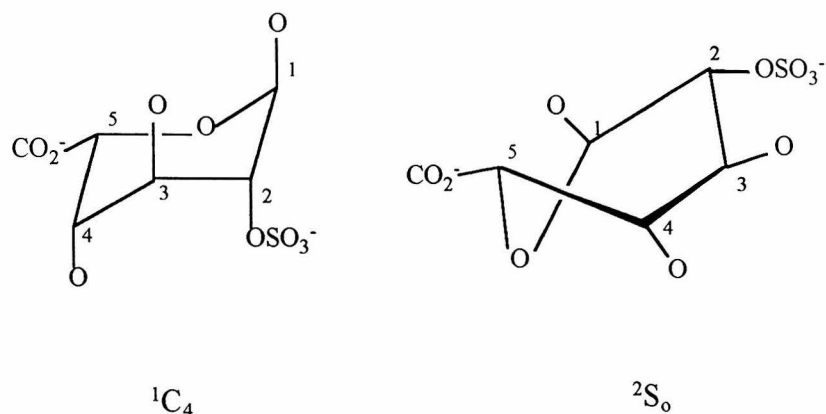


Fig. 2.5. The 1C_4 chair and 2S_0 skew boat ring conformations of L-iduronic acid. The nomenclature describes the conformation of the six membered hexopyranose ring (C or S for chair and skew boat, respectively), while the super- and sub-scripted numbers refer to the two ring atoms most displaced from the plane defined by the other four ring atoms (14).

Discussion

The structures of the heparin tetrasaccharide and hexasaccharide cocrystallized with human bFGF directly demonstrate that the primary binding site for heparin on this protein is generated by residues 27-28, 101-103, and 120-137. With the exception of the region around Asn-102, these assignments are in excellent agreement with previously identified bFGF residues associated with heparin binding. Comparisons of the binding interactions exhibited between bFGF and the tetrasaccharide and hexasaccharide fragments indicates, at least for the sequences used in this study, that the heparin binding region may be further separated into higher and lower affinity regions. The higher

affinity binding site interacts with heparin through the sidechains of residues Asn-28, Arg-121, Lys-126 and Gln-135, while the lower affinity site utilizes the sidechains of residues Lys-27, Asn-102, and Lys-136. The four residues that form the higher affinity region correspond exactly to the four residues identified by Thompson, et al. (15), as having the greatest impact on the binding of bFGF to heparin when mutated to alanine. Conversely, residue substitutions that had relatively little effect on heparin binding in that study correspond to residues found either in the lower affinity binding region, or which do not appear to be directly involved in heparin binding.

Many of the specific interactions involve multiple interactions between sites on bFGF and negatively charged groups from the heparin fragments. The most extensive interactions in both the higher and lower affinity sites are provided by the 2-*N*-sulfate groups of glucosamine and the 2-*O*-sulfate group of iduronic acid; the 6-carboxyl group of iduronic acid also participates in a salt bridge contact found in the lower affinity site (Table 2.1). These observations are consistent with previous studies on the interactions between bFGF and specifically desulfated heparin analogs (16). No contacts between bFGF and the 6-*O*-sulfate groups of glucosamine are observed, as these groups are directed away from the bFGF surface, and are potentially available to interact with other ligands, perhaps including FGF receptors.

Comparison with other anion-bound FGF structures

Previous crystallographic studies have characterized the binding modes of other anionic compounds, including sulfate ions and sucrose octasulfate, to members of the FGF family. The superposition of these structures onto the heparin fragment-bFGF complexes demonstrates that while these compounds all bind to the same general region

of the FGF surface, due to their distinct chemical structures, the specific interactions are different (Fig. 2.6). The greatest similarity in these structures comes from the position of one particular sulfate that participates in crucial recognition contacts on the FGF surface. This sulfate, which corresponds to the Glcn2 2-*N*-sulfate of the heparin fragment, corresponds to a sulfate group of sucrose octasulfate (11) and a bound sulfate from crystallization of the unliganded protein (17). This sulfate position binds to the higher affinity binding region of bFGF, and interacts with the sidechains of Asn-28 and Lys-126. The other seven sulfates of sucrose octasulfate, while binding to the same general region of the FGF surface, do not directly correspond to other specific anionic groups of the heparin fragments. Likewise, two bound sulfate groups from the solution used to crystallize the unliganded protein also do not coincide with specific groups of the heparin fragments; one of these sulfates binds to a position corresponding to the heparin hexasaccharide, but does not coincide with any of the heparin sulfates, while the third sulfate interacts with Lys-120 and Lys-130, away from the heparin fragment in the hexamer-bFGF structure. Comparison of the hexamer-bFGF structure to that of a bFGF

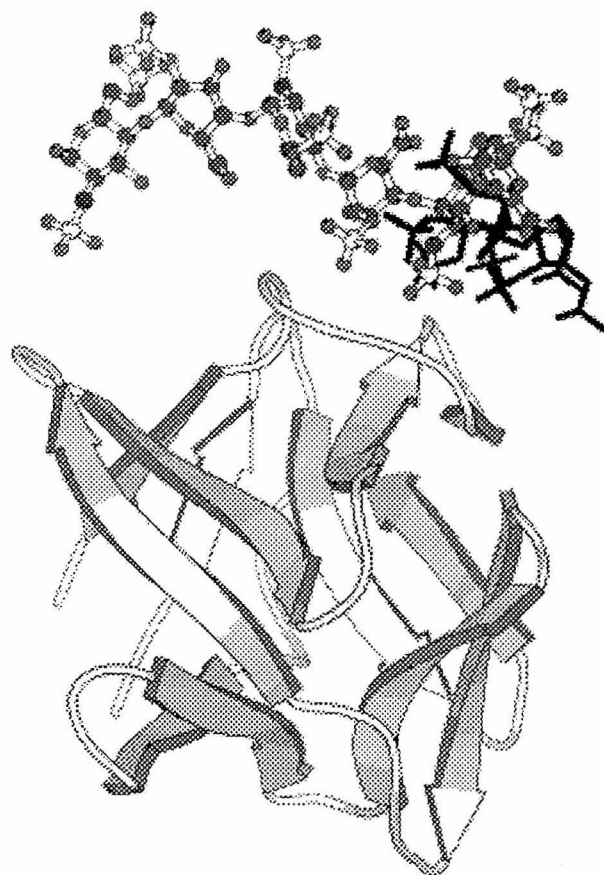


Fig. 2.6. Superposition of the aFGF-sucrose octasulfate and bFGF-heparin hexasaccharide structures.

complexed with a desulfated heparin trisaccharide analog revealed that while one of the two observed trisaccharide is bound at a remote site, the other is bound at the heparin binding region. Analysis of these two structures shows little similarity in their specific interactions, due to the desulfation of the trisaccharide analog. Consequently, while many anions bind to this general region of bFGF, the detailed binding interactions can vary appreciably.

Sequence alignment

The FGF family is currently composed of 12 members (18), with sequence identities varying from 30-55%. Sequence alignment for nine of its members was performed (FGF1-FGF9), and the conservation of the residues interacting with heparin was evaluated. No residue in the heparin binding region is completely conserved throughout the FGF family; furthermore, even the charge of the residue sidechain fails to be strictly maintained. As a result of these sequence variations, it is likely that different members of the FGF family will utilize different contacts for the binding to heparin. Although bFGF makes a specific contact to the 2-*O*-sulfate group of Idu through the sidechain of Gln-135, the substitution of this residue by Met in HST should interfere with this interaction, consistent with the observation that HST can bind heparin in the absence of 2-*O*-sulfate groups (16). The replacement of Lys-27 and Asn-102 in bFGF by Ala and Leu in INT2 should prevent interaction of these residues with the GlcN 2-*N*-sulfate group in the lower affinity site. Perhaps the most dramatic substitutions involve Asn-28 and Lys-126, which are changed to Thr and Val, respectively, in KGF. These substitutions would be expected to block contact of these sidechains with the GlcN 2-*N*-sulfate group in the high affinity site. This alteration in the interaction between heparin and KGF

relative to bFGF may contribute to the differential effects of heparan sulfates on KGF relative to aFGF (19).

These findings complement previous reports that FGF binding and signaling requirements were different for the different growth factors (16). bFGF requires 2 O-sulfate for binding, but not the 6 O-sulfate; on the other hand aFGF requires both the 2 O-sulfate and the 6 O-sulfate for heparin binding, and HST(FGF4) can bind heparin in the absence of either 6 O-sulfate or the 2 O-sulfate.

Models on FGF activation

Comparison of the unliganded bFGF structure to the structures of bFGF complexed with heparin fragments or other anionic molecules demonstrates that complex formation is not accompanied by significant change in the polypeptide conformation of bFGF. Consequently, it appears unlikely that a heparin induced conformational change in bFGF mediates initiation of the signal transduction pathway. The present structural studies instead support the proposal that the role of heparin is to either bind multiple FGF molecules to promote receptor dimerization or to facilitate formation of an FGF - receptor complex by binding both components (16,20,21,22). The proposal that protein-protein associations involved in signal transduction pathways may be regulated by binding interactions to specific molecules such as heparin is not unique to the FGF system, but reflects a mechanism of general biological relevance (23).

Since heparin fragments containing 12-14 sugar residues are required for stimulation of FGF activity, the tetra- and hexasaccharide structures cannot directly address details of the interactions between these longer heparin fragments with one or more FGF molecules. However, some general aspects of this interaction can be

suggested from the structures of these shorter heparin fragments. The heparin tetra- and hexasaccharides are consistent to a good first approximation with a helical structure having a 17Å repeat and containing 4 sugar residues. These 4 residues may be further considered to be generated from a disaccharide repeat where consecutive disaccharides face out from opposite sides of the heparin helix. The diameter of an FGF molecule in the direction of an extended heparin molecule may be estimated as ~33Å. Consequently, a single FGF molecule binding to heparin would occupy the equivalent of ~8 sugars. Two limiting modes of interaction between a heparin polymer and multiple FGF molecules may be envisioned: for molecules such as bFGF which bind to heparin with a limiting density of ~1 bFGF per 8 sugar residues (24), it is possible that the FGF molecules are translationally ordered along one side of the heparin helix. Alternatively, in the case of aFGF which binds to heparin with a limiting density of ~1 aFGF per 4-5 sugar residues (25), it is possible that the aFGF molecules are distributed along both sides of the heparin helix. In either model, it is possible that binding sites along the heparin molecule would be available for interaction with additional ligands such as FGF receptors; these molecules could interact with heparin groups that are otherwise not utilized in FGF binding, such as the 6-*O*-sulfate groups, or they could bind to sites on the heparin not occupied by FGF and interact with the same groups. It is possible that the precise sequence of sugar residues in heparin might be used to specify the presence or absence of FGF binding sites to facilitate the appropriate assembly of molecules required for initiation of signal transduction. With the sequence variability exhibited by both heparin and different members of the FGF family, it is likely that a diverse set of binding interactions between FGFs and heparin-like molecules occur, and these may be reflected in the differing responses and requirements for these sulfated polysaccharides exhibited by the FGF family.

Drug design

It has been reported that a minimum sequence of five sugar units is required for binding to bFGF, but it is clearly shown in this structure that a tetrasaccharide is sufficient. Actually, it could be inferred from this structure due to the lack of contacts made by rings 1 and 4 that a two sugar units molecule would be capable of bFGF binding as long they are locked in the proper conformation, which seems to be the main function of the terminal sugar units in the tetrasaccharide case. However, that may not be true for the other FGF family members, as it has been demonstrated that binding requirements are different for the different members.

The results presented here can be valuable for future drug design experiments. The structures of the various complex ligands could be used to set guidelines for the design of molecules that can bind FGF. For example, the conserved sulfate position can be used as a handle onto which the various ligands can be modeled. Clearly, FGFs can accommodate different types of ligands as has often been demonstrated. Solving more structures of complexes between FGFs and the various ligands might allow us to understand the codes behind ligand binding and specificity.

heparin		bFGF		closest distance (Å)	
residue #	atom (or group)	residue	atom	tetramer	hexamer
GlcN-2	2-SO ₃ ²⁻	N28	Nδ2	3.3	3.2
GlcN-2	2-SO ₃ ²⁻	R121	NH	3.1	3.1
GlcN-2	2-SO ₃ ²⁻	K126	Nζ	3.0	2.9
IdU-3	3-OH	N28	Oδ1	2.8	2.6
IdU-3	3-OH	R121	Nη1	3.2	3.1
IdU-3	2-O-SO ₃ ²⁻	K126	Nζ	2.8	3.0
IdU-3	2-O-SO ₃ ²⁻	Q135	Nε2	3.5	3.4
IdU-3	2-O-SO ₃ ²⁻	K136	NH	3.0	3.1
IdU-3	2-O-SO ₃ ²⁻	A137	NH	2.9	3.0
IdU-5	6-CO ₂ ⁻	K136	Nζ		3.3
GlcN-6	2-SO ₃ ²⁻	K27	Nζ		3.6
GlcN-6	2-SO ₃ ²⁻	N102	Nδ2		2.5
GlcN-6	3-OH	K27	Nζ		2.8

Table 2.1. Polar (ion pair and hydrogen bond) contacts between bFGF residues and tetrasaccharide and hexasaccharide fragments of heparin. Interactions between atoms separated by less than 3.6 Å are indicated.

glycosidic bond	Φ		Ψ	
	x-ray	NMR	x-ray	NMR
(ring 1-2)	57°		8°	
(ring 2-3)	-27°	-39°	-3°	-33°
(ring 3-4)	50°	41°	24°	14°
(ring 4-5)	40°	-9°	-41°	-41°
(ring 5-6)	48°	61°	2°	16°

Table 2.2. Values of the ϕ , ψ torsion angles about the glycosidic bonds in the hexasaccharide fragment of heparin complexed to bFGF. ϕ and ψ are defined as the torsion angles about the H1-C1-O4'-C4' and the C1-O4'-C4'-H4', respectively, that connect successive sugar rings by 1,4 linkages. NMR-derived values are from (13).

member	bFGF residue number						
	27	28	102	121	126	135	136
bFGF	K	N	N	R	K	Q	K
aFGF	S	N	N	K	K	Q	K
INT2	A	T	L	G	R	Q	K
HST	N	V	N	K	K	M	K
FGF5	R	V	N	K	K	H	I
FGF6	N	V	N	K	K	M	T
KGF	R	T	N	Q	V	Q	K
AIGF	R	T	N	R	R	Q	R
FGF9	R	T	N	K	R	Q	K
heparin	GlcN2-SO ₃ ²⁻	GlcN2-SO ₃ ²⁻	GlcN2-SO ₃ ²⁻	Idu3-OH	GlcN2-SO ₃ ²⁻	Idu2-OSO ₃ ²⁻	Idu6-CO ₂ ⁻
contact	GlcN3-OH	Idu3-OH			Idu2-OSO ₃ ²⁻		
affinity	low	high	low	high	high	high	low

Table 2.3. Sequence alignment of members of the FGF family in the heparin binding region. The GCG package was used to generate the alignment. Heparin residues contacting side chain atoms of the indicated residues are listed, with the abbreviations G = glucosamine and I = iduronic acid. High affinity refers to contacts involving GlcN-2 or Idu-3, while low affinity refers to contacts involving Idu-5 or GlcN-6. The accession numbers for the bFGF, aFGF, INT2, HST, FGF5, FGF6, KGF, AIGF and FGF9 sequences are A26642, A24820, S04742, Tvhuhs, Tvhuf5, S04204, A36301, A46245, and A48137, respectively, of the PIR database.

References

- (1) A. Pervin, C. Gallo, K. A. Jandik, X.-J. Han, R. J. Linhardt, *Glycobiology* **5**, 83-95 (1995).
- (2) G. M. Fox, et al., *J. Biol. Chem.* **263**, 18452-18458 (1988).
- (3) S. Bailey, *Acta Cryst D* **50**, 760-763 (1994).
- (4) W. Kabsch, *J. Appl. Cryst.* **21**, 916-924 (1988).
- (5) J. Zhang, L. S. Cousens, P. J. Barr, S. R. Sprang, *Proc. Natl. Acad. Sci. USA* **88**, 3446-3450 (1991).
- (6) A. T. Bruger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
- (7) T. A. Jones, *Methods Enzymol* **115**, 157 (1985).
- (8) W. Kabsch, *Acta crystallogr.* **A32**, 922-923 (1976).
- (9) J. M. Cox, *J. Mol. Biol.* **28**, 151-156 (1967).
- (10) P. J. Kraulis, *J. Appl. Cryst.* **24**, 946-950 (1991).
- (11) X. Zhu, B. T. Hsu, D. C. Rees, *Structure* **1**, 27-34 (1993).
- (12) D. M. Ornitz, et al., *Science* **268**, 432-436 (1995).
- (13) B. Mulloy, M. J. Forster, C. Jones, D. B. Davies, *Biochem. J.* **293**, 849-858 (1993).
- (14) J. J. Stoddart, *Stereochemistry of carbohydrates* (Wiley-Interscience, New York, 1971).
- (15) L. D. Thompson, M. W. Pantoliano, B. A. Springer, *Biochemistry* **33**, 3831-3840 (1994).
- (16) S. Guimond, M. Maccarana, B. B. Olwin, U. Lindahl, A. C. Rapraeger, *J. Biol. Chem.* **268**, 23906-23914 (1993).
- (17) J. Zhang, L. S. Cousens, P. J. Barr, S. R. Sprang, *Proc. Natl. Acad. Sci. USA* **88**, 3446-3450 (1991).

- (18) A. S. Verdier, M. G. Mattei, H. Lovec, H. Harung, M. Goldfarb, D. Birnbaum, *Genomics* **40**, 151-154 (1997).
- (19) R. Reichslotky, et al., *J. Biol. Chem.* **269**, 32279-32285 (1994).
- (20) D. M. Ornitz, et al., *Mol. Cell. Biol.* **12**, 240-247 (1992).
- (21) T. Spivak-Kroizman, et al., *Cell* **79**, 1015-1024 (1994).
- (22) M. Kan, et al., *Science* **259**, 1918-1921 (1993).
- (23) D. J. Austin, G. R. Crabtree, S. L. Schreiber, *Chem. & Biol.* **1**, 131-136 (1994).
- (24) T. Arakawa, J. Wen, J. S. Philo, *Arch. Biochem. Biophys.* **308**, 267-273 (1994).
- (25) H. Mach, et al., *Biochem.* **32**, 5480-5489 (1993).

Supplemental material

Following are the generate, parameter, and topology files used for refinement in X-PLOR. Every sugar unit in the heparin fragment is defined as separate residue.

remarks file GENERATE.INP

remarks generate structure file and hydrogens for test case

topology @../toph19x.pro

@toph19x.hep6

@../toph11.wat

end { * read topology file * }

parameter

@../prolsq.pro { * read empirical potential * }

@prolsq.hep6

@../prolsq.wat

{ * with modifications * }

nbonds { * this statement specifies * }

atom cdie shift eps=1.0 e14fac=0.4 { * the nonbonded interaction * }

cutnb=7.5 ctonnb=6.0 ctofnb=6.5 { * energy options. Note the * }

nbxmod=5 vswitch { * reduced nonbonding cutoff * }

end { * to save some CPU time * }

end

segment { * generate protein * }

name=" " { * this name has to match the four characters in columns*}
{ * 73 through 76 in the coordinate file, in XPLOR this *}
{ * name is referred to as SEGId. * }

chain

@../toph19.pep { * read peptide bond file * }

coordinates @modelhep6_p.pdb { * interpret coordinate file to obtain sequence * }

end

end

vector do (name="O") (name OT1) { * sometimes different * }

vector do (name="OT") (name OT2) { * atom names are used, * }

vector do (name="CD1") (name CD and resname ile) { * to rename them, un- * }
{ * comment these lines * }

coordinates @modelhep6_p.pdb { * here we actually read the coordinates * }

segment { * generate protein * }

name=" " { * this name has to match the four characters in columns*}
{ * 73 through 76 in the coordinate file, in XPLOR this *}
{ * name is referred to as SEGId. * }

chain

{ * read peptide bond file * }

```

coordinates @modelhep6_h.pdb { * interpret coordinate file to obtain sequence * }
end
end
vector do (name="O") ( name OT1 )           { * sometimes different * }
vector do (name="OT") ( name OT2 )          { * atom names are used, * }
vector do (name="CD1") ( name CD and resname ile ) { * to rename them, un- * }
                                         { * comment these lines * }

```

```

coordinates @modelhep6_h.pdb      { * here we actually read the coordinates * }

```

```

segment                          { * generate protein * }

```

```

name=" "      { * this name has to match the four characters in columns* }
              { * 73 through 76 in the coordinate file, in XPLOR this * }
              { * name is referred to as SEGIId.                * }
chain
              { * read peptide bond file * }
coordinates @modelhep6_w.pdb { * interpret coordinate file to obtain sequence * }
end
end
vector do (name="O") ( name OT1 )           { * sometimes different * }
vector do (name="OT") ( name OT2 )          { * atom names are used, * }
vector do (name="CD1") ( name CD and resname ile ) { * to rename them, un- * }
                                         { * comment these lines * }

coordinates @modelhep6_w.pdb      { * here we actually read the coordinates * }

```

```

{ * the generation of co-factors, waters,... would * }
{ * follow here. Note, that one has to split the * }
{ * coordinate file into separate files containing * }
{ * the protein, co-factors, substrate, water,... * }

```

```

{ * the following shows how to create a S-S bridge * }

```

```

patch L14
reference=1=( resid 301 )
reference=2=( resid 302 )
end

```

```

patch L41
reference=1=( resid 302 )
reference=2=( resid 303 )
end

```

```

patch L14
reference=1=( resid 303 )
reference=2=( resid 304 )
end

```

```

patch L41
reference=1=( resid 304 )
reference=2=( resid 305 )
end

```

```

patch L14
  reference=1=( resid 305 )
  reference=2=( resid 306 )
end

flags exclude vdw elec end    { * do hydrogen building w/o vdw and elec  *}

hbuild                      { * this statement builds missing hydrogens *}
  selection=( hydrogen )    { * which are needed for the force field  *}
  phistep=4
end

write coordinates output=genhep.pdb end    { * write out coordinates *}

write structure output=genhep.psf end { * this writes the structure file *}
stop

```

```

REMARKS prolsq.hep6
REMARKS =====
REMARKS This dictionary is modified from param1.cho. Dramatic change is
REMARKS made for the energy term by using prolsq.pro.

```

```

REmarks attention: Some of this is duplicate of prolsq.pro. This can
REmarks          be trouble running with proteins.

```

```

remark - parameter file PROLSQ.PRO -
remark Trying to match PROLSQ energy functions - ak

```

```

set echo=false end

```

```

evaluate ( $kbond    = 2500.0 ) !dimension: ( kcal / mole-A^2 )
evaluate ( $kangle   = 584.0 ) !dimension: ( kcal / mole-rad^2 )
evaluate ( $omega    = 1700.0 ) !dimension: ( kcal / mole    )
evaluate ( $pro_omega = 100.0 ) !dimension: ( kcal / mole    )
evaluate ( $plane    = 360.0 ) !dimension: ( kcal / mole-rad^2 )
evaluate ( $chiral   = 796.0 ) !dimension: ( kcal / mole-rad^2 )

```

```

bond CC  CC  $kbond 1.523
bond CC  OE  $kbond 1.427
bond CC  NH1 $kbond 1.45
bond CC  OH1 $kbond 1.411
bond CC  C   $kbond 1.52
bond S   OE  $kbond 1.77
bond CC  CD  $kbond 1.44
bond CD  CD  $kbond 1.33
bond CD  C   $kbond 1.33
bond CD  OE  $kbond 1.34

```

bond NH1 S \$kbond 1.77

angle CC CC CC \$kangle 110.7
angle OE S OC \$kangle 109.5
angle CC CD CD \$kangle 120.0
angle CD CD C \$kangle 120.0
angle CD CD OE \$kangle 120.0
angle CC CC CD \$kangle 109.5
angle C CD OE \$kangle 120.0
angle CC OE CD \$kangle 115.0
angle CD CC OH1 \$kangle 109.5
angle CD C OC \$kangle 120.0
angle CC OE S \$kangle 119.9
angle OH1 CC OE \$kangle 111.55
angle CC NH1 S \$kangle 119.9
angle NH1 S OC \$kangle 109.5
angle CC CC OE \$kangle 109.4
angle CC OE CC \$kangle 113.8
angle OE CC OE \$kangle 110.7
angle OE CC C \$kangle 107.24
angle CC CC OH1 \$kangle 110.1
angle CC CC NH1 \$kangle 110.0
angle CC CC C \$kangle 108.7
angle CC C OC \$kangle 118.7

dihedral X CC CC X 50.00 3 0.0
dihedral X CC OH1 X 50.00 3 0.0
dihedral X CC OE X 50.00 3 0.0

{ These high force constants maintain chiral centers on the rings }

impr CD CD C OE 500.0 0 0.0
impr C CD OC OC 500.0 0 0.0
impr OE CD C OC 500.0 0 0.0
impr CC CD CD OE 500.0 0 0.0
impr CD CD C OC 500.0 0 0.0
impr C OC OC CC 100.0 0 0.0

{ This is as other acids }

!!for use with:

!!NBXMOD=5 ATOM CDIEL SHIFT vswitch

!! CUTNB=8.0 CTOFNB=7.5 CTONNB=6.5 EPS=1.0 E14FAC=0.4 WMIN=1.5

NONBonded OE 0.1591 2.8509 0.1591 2.8509
NONBonded CD 0.1000 3.2963 0.1000 3.0291
NONBonded CC 0.0903 3.2072 0.0903 3.2072
NONBonded HH 0.1000 2.1381 0.1000 1.8709

set echo=true end

REMARKS toph19x.hep6
 REMARKS =====
 REMARKS modified by SF (APR/1995)
 REMARKS This pro file is used with toph19x.pro (for proteins).
 REMARKS So the overlapping part with toph19x.pro is taken out.

set echo=false end

AUTOGENERATE ANGLES=TRUE END
 {*=====*

{default masses -- only those not in protein topology TOPH19X.PRO}

MASS CC 12.01100! carbohydrate carbon
 MASS OE 15.99940! ether oxygen
 MASS HA 1.00800! aliphatic hydrogen
 MASS CD 12.01100!

!RESIdue HEP ! Heparin
 RESIdue GCN ! Glucosamine
 GROUP

!mono

ATOM C1 TYPE=CC CHARGE= 0.000 END
 ATOM C2 TYPE=CC CHARGE= 0.000 END
 ATOM C3 TYPE=CC CHARGE= 0.000 END
 ATOM C4 TYPE=CC CHARGE= 0.000 END
 ATOM C5 TYPE=CC CHARGE= 0.000 END
 ATOM C6 TYPE=CC CHARGE= 0.000 END
 ATOM O1 TYPE=OE CHARGE= 0.000 END
 ATOM O3 TYPE=OH1 CHARGE=-0.400 END
 ATOM O4 TYPE=OE CHARGE= 0.000 END
 ATOM O5 TYPE=OE CHARGE= 0.000 END
 ATOM O6 TYPE=OE CHARGE=-0.020 END
 ATOM O22 TYPE=OC CHARGE=-0.660 END
 ATOM O23 TYPE=OC CHARGE=-0.660 END
 ATOM O24 TYPE=OC CHARGE=-0.660 END
 ATOM O62 TYPE=OC CHARGE=-0.660 END
 ATOM O63 TYPE=OC CHARGE=-0.660 END
 ATOM O64 TYPE=OC CHARGE=-0.660 END
 ATOM N2 TYPE=NH1 CHARGE=-0.020 END
 ATOM S2 TYPE=S CHARGE= 0.000 END
 ATOM S6 TYPE=S CHARGE= 0.000 END

!mono

BOND C1 O1
 BOND C1 O5
 BOND C1 C2
 BOND C2 N2
 BOND C2 C3
 BOND C3 O3
 BOND C3 C4
 BOND C4 O4

BOND C4 C5
 BOND C5 O5
 BOND C5 C6
 BOND C6 O6
 BOND N2 S2
 BOND S2 O22
 BOND S2 O23
 BOND S2 O24
 BOND O6 S6
 BOND S6 O62
 BOND S6 O63
 BOND S6 O64

DIHEDRAL C1 C2 C3 C4
 DIHEDRAL C2 C3 C4 C5
 DIHEDRAL C3 C4 C5 O5
 DIHEDRAL C4 C5 O5 C1
 DIHEDRAL C5 O5 C1 C2
 DIHEDRAL O5 C1 C2 C3

END

!di

RESIdue UAP

GROUP

ATOM C1 TYPE=CC CHARge= 0.000 END
 ATOM C2 TYPE=CC CHARge= 0.000 END
 ATOM C3 TYPE=CC CHARge= 0.000 END
 ATOM C4 TYPE=CD CHARge= 0.000 END
 ATOM C5 TYPE=CD CHARge= 0.000 END
 ATOM C6 TYPE=C CHARge= 0.000 END
 ATOM O2 TYPE=OE CHARge= 0.000 END
 ATOM O22 TYPE=OC CHARge= 0.000 END
 ATOM O23 TYPE=OC CHARge= 0.000 END
 ATOM O24 TYPE=OC CHARge= 0.000 END
 ATOM O3 TYPE=OH1 CHARge= 0.000 END
 ATOM O5 TYPE=OE CHARge= 0.000 END
 ATOM O6 TYPE=OC CHARge= 0.000 END
 ATOM O7 TYPE=OC CHARge= 0.000 END
 ATOM S2 TYPE=S CHARge= 0.000 END

!di

BOND C1 C2
 BOND C1 O5
 BOND C2 C3
 BOND C2 O2
 BOND C3 C4
 BOND C3 O3
 BOND C4 C5
 BOND C5 C6
 BOND C5 O5
 BOND C6 O6
 BOND C6 O7
 BOND O2 S2

BOND S2 O22
BOND S2 O23
BOND S2 O24

! improper terms for c=c and carboxylate groups

IMPRoper C5 C4 C6 O5
IMPRoper C6 C5 O6 O7
IMPRoper C4 C5 C6 O6
IMPRoper C3 C4 C5 O5
IMPRoper O5 C5 C6 O7

END

RESIdue IDU !Iduronic acid
GROUP

!ring 3

ATOM C1 TYPE=CC CHARGE= 0.000 END
ATOM C2 TYPE=CC CHARGE= 0.000 END
ATOM C3 TYPE=CC CHARGE= 0.000 END
ATOM C4 TYPE=CC CHARGE= 0.000 END
ATOM C5 TYPE=CC CHARGE= 0.000 END
ATOM C6 TYPE=C CHARGE= 0.000 END
ATOM O2 TYPE=OE CHARGE= 0.000 END
ATOM O22 TYPE=OC CHARGE= 0.000 END
ATOM O23 TYPE=OC CHARGE= 0.000 END
ATOM O24 TYPE=OC CHARGE= 0.000 END
ATOM O3 TYPE=OH1 CHARGE= 0.000 END
ATOM O5 TYPE=OE CHARGE= 0.000 END
ATOM O6 TYPE=OC CHARGE= 0.000 END
ATOM O7 TYPE=OC CHARGE= 0.000 END
ATOM S2 TYPE=S CHARGE= 0.000 END

!RING 3

BOND C1 C2
BOND C1 O5
BOND C2 C3
BOND C2 O2
BOND C3 C4
BOND C3 O3
BOND C4 C5
BOND C5 C6
BOND C5 O5
BOND C6 O6
BOND C6 O7
BOND O2 S2
BOND S2 O22
BOND S2 O23
BOND S2 O24

DIHEDRAL C1 C2 C3 C4

DIHEDRAL C2 C3 C4 C5
DIHEDRAL C3 C4 C5 O5
DIHEDRAL C4 C5 O5 C1
DIHEDRAL C5 O5 C1 C2
DIHEDRAL O5 C1 C2 C3

IMPRoper C6 O6 O7 C5

END

!Link sugar units!

PRESide L14

add bond 1C1 2O4

add angle 1C1 2O4 2C4

add angle 1C2 1C1 2O4

Add angle 1O5 1C1 2O4

END

PRESide L41

add bond 1O1 2C4

add angle 1C1 1O1 2C4

add angle 1O1 2C4 2C3

add angle 1O1 2C4 2C5

END

Chapter 3

Crystal Structure of Keratinocyte Growth Factor

Experimental Procedures

Human KGF with a deletion of the first 24 residues ($\Delta 24$) was constructed and overexpressed in *E. coli* (kindly provided by Amgen). Crystals of KGF were grown using the vapor diffusion method, using 1.0 μ l of 4% inositol hexaphosphate (IHP), 3.0 μ l of a 10.0 mg/ml KGF solution, and 3.0 of reservoir solution. The reservoir solution contained 10% PEG 1K, and 0.1M Tris pH 7.5. Crystals grew over a few days at room temperature. These crystal were characterized as belonging to space group C2 and diffracted to ~ 3.5 Å resolution. Upon addition of 2.0 μ l of 50% ethylene glycol solution, a new crystal form grows out of the solution in a few days. This crystal form was characterized as belonging to the space group $P2_12_12_1$ with cell dimensions of $a=33.4$ Å, $b=68.0$ Å, $c=124.0$ Å. Diffraction data was collected on a Siemens multiwire area detector at room temperature. Data was processed using the program XDS (1). The data set contained 10821 reflections to 2.4 Å (91.8% complete), merged from 75031 observations with a merging R of 0.048. Two molecules were present in the asymmetric unit, and native Patterson maps indicated that they were related by (0.36, 0.22, 0.50) translation, due to the presence of a packing peak at that site (Fig. 3.1). The packing peak indicated the presence of a local non-crystallographic twofold axis parallel to the Z axis.

Molecular replacement (MR) was performed using human bFGF as a starting model (2) for both crystal forms using AMORE (3). For the orthorhombic crystal form, the first molecule was located by a typical (MR) run in Amore, and the second was generated by taking advantage of the packing peak present in the native Patterson map.

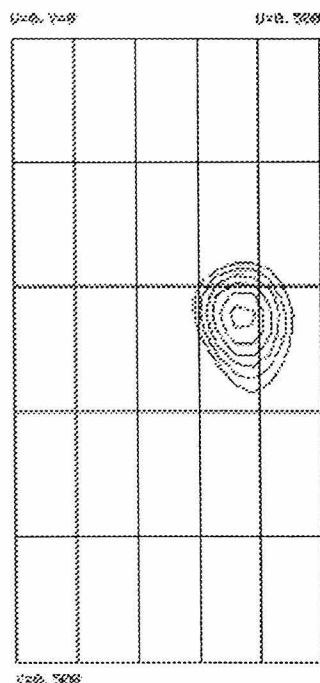


Fig. 3.1. Harker section $W=1/2$ of the native Patterson map of the KGF crystals showing the packing peak at (0.36,0.22,0.5). The Patterson map was calculated using data between 20.0 to 6.0 Å resolution.

After rigid body refinement in X-PLOR (4), the model was truncated to a polyalanine model. Solvent flattening and twofold averaging were then performed in SOLOMON (CCP4) (5). The map at that point was of a good quality allowing model building with high confidence. Model building was performed using the TOM/FRODO graphics program (6). Traditional positional refinement and slowcool were performed in XPLOR. Non-crystallographic symmetry restraints were used at intermediate stages of the

refinement. The structure refined to $R=0.21$ and $R_{\text{free}}=0.32$ for data between 8.0 and 2.4 Å resolution, with $F > \sigma(F)$, and RMS deviations for bond distances and angles from ideal values of 0.013 Å and 2.37° respectively. Many of the side chains are disordered (up to 18), as well as the IHP. Superpositions of the different structures were based on the backbone C_{α} positions and were done using a program based on the algorithm of Kabsch (7). Structural figures were prepared with the program MOLSCRIPT (8).

Results and Discussion

Overall Structure

The overall fold of KGF is similar to that of other FGF protein family members (2), and consists of a β barrel fold with 12 β strands (Fig. 3.2).

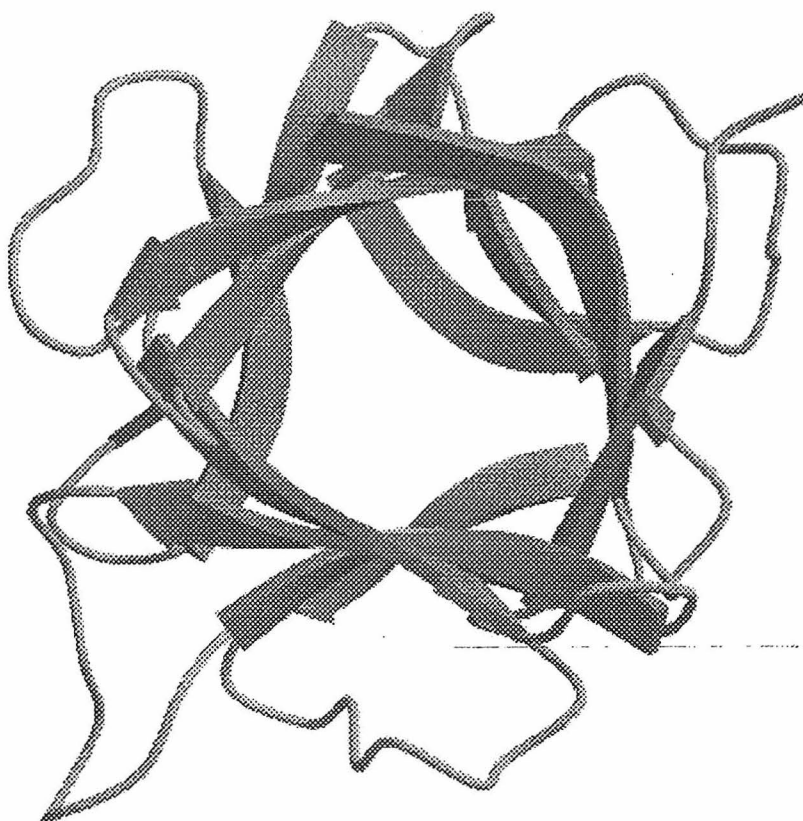


Fig. 3.2. Ribbons diagram of the three-dimensional structure of KGF.

When compared to the structure of bFGF, the KGF structure reveals three extra residues ordered at the amino terminus, and two more ordered residues at the carboxy terminus. The extra three residues at the N-terminus adopt a β strand conformation extending the β sheet formed by strands 4, 5, 6, and 7, and making the first β strand unusually bent. This is unlike the loop conformation of the ordered aFGF N-terminus. The N-termini of aFGF and bFGF do not contain a typical secretion sequence, and thus do not have the same function as those of the other FGF members. The extra two residues at the C-terminus adopt a loop conformation pointing away from the protein. The overall RMS deviation between KGF and bFGF based on 122 of the C_{α} backbone atoms is 1.07 Å, and that between KGF and aFGF is 1.33 Å based on 123 C_{α} backbone atoms (Fig. 3.3). The RMS deviation between the two KGF molecules is 0.4 Å.

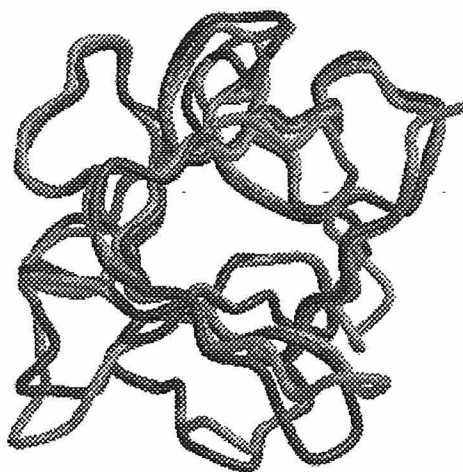


Fig. 3.3. Superposition of KGF with aFGF. KGF is in light gray.

Disulfide bond

Another new feature of the KGF structure is a disulfide bond that is formed between cysteines 102 and 106. These cysteines are separated by only 4 residues, and are both present in loop 7. This disulfide bond might help stabilize the native loop conformation, increasing the stability of the native protein structure. It has actually been shown that substitution of all cysteines to alanine or serine in KGF leads to destabilization of its structure, and its inability to bind heparin any longer. This is likely to be a result of losing the entropic stability provided by the disulfide bond. Cys106 is conserved in all nine members of the FGF family. However, Cys102 is unique to KGF, and is replaced by a variety of amino acids in other FGF members, indicating that this disulfide bond is unique to KGF.

Structural comparison and functional implications

Heparin

A comparison of the KGF structure with the crystal structure of bFGF complexed with the heparin hexamer fragment (chapter 2) (9) has been performed. The bFGF-heparin hexamer structure revealed two heparin binding sites, a high affinity site and low affinity one. Upon superpositioning the bFGF-Heparin complex with KGF, a conceivable KGF heparin binding site is revealed. Five positively charged residues are found in close proximity to the heparin molecule, thus suggesting that the heparin binding site is maintained; however, the specific interactions are different. Arg27, His102, and Lys139 of KGF are found in the low affinity site, while Arg130 and Lys133 are found in the high

affinity site. This implies that the low affinity site in bFGF might be a high affinity site in KGF, and vice versa. The neutral residues Asn102 and Gln135 make contacts with heparin in the bFGF-heparin complex, and are conserved in KGF; thus they can form similar contacts with heparin. It has been suggested that the heparin binding mechanism with the various FGFs is variable (10).

Loop regions

When the KGF structure is compared to both aFGF, and bFGF, most of the differences appear in the loop regions. Out of the 11 loops separating the 12 β strands, 4 appear to have significant differences from both aFGF and bFGF, as defined by more than 2.0 Å RMS deviation in the C_α positions upon superimposing the structures. These are loops (1, 6, 8, and 9). In loop 1, KGF is missing a residue. The shorter loop is accommodated by a tighter turn in KGF. Loop 6 has a significantly different conformation in KGF, despite having the same size as that of bFGF. Differences are also found in the β strand following loop 6, leading to the only strand with a different orientation. Loop 8, despite having the same size as in bFGF and an identical sequence to aFGF, has a different conformation, and in this crystal form is involved in interprotein contacts. Loop 9 is the site of a four-residue insertion in KGF, when compared to bFGF. These extra residues make the loop longer, maintaining some backbone hydrogen bonding through turn, indicating that a longer insertion, as in INT2, might lead to a β turn β fold.

The structural differences between KGF and bFGF are likely to be responsible for their functional differences. Since the major backbone conformational differences are localized in four loops, it is possible to attribute receptor binding specificity to these regions, in analogy to what has been shown in the IL8 case (11). These four loops are in two regions (I and II) on one side of the protein (Fig. 3.4). These two regions cover a similar surface as the two regions identified by a mutagenesis study to be important for receptor binding.

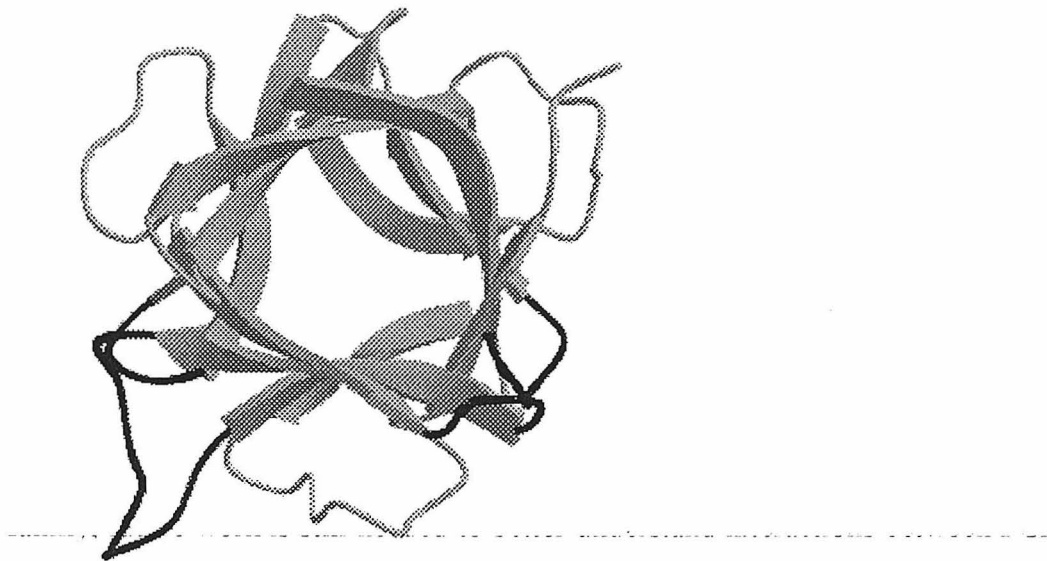


Fig. 3.4. Ribbons diagram of KGF showing in black the loops that have the largest structural differences from aFGF and bFGF.

Interestingly, the heparin binding site falls on the same side of the protein, and is in the middle, separating these two regions, thus indicating that heparin and the receptor bind on the same face of the protein. This type of binding mechanism allows for direct interactions between the receptor and heparin, which has been proposed (Fig. 3.5).

Based on sequence alignment of nine members of the FGF family (FGF1-FGF9), size dependent conformational changes are expected for loops 3, 4, 7, and 11. These loops are positioned on the two sides of the protein mentioned earlier; however, they are farther away from the heparin binding site (Fig. 3.6). These loops are sufficiently far apart that one Ig domain could not cover the entire area, so that a second Ig domain might come into play. As mentioned earlier, there is evidence that KGF is recognized by only one Ig domain, the Ig3 domain. However, other FGF members are believed to interact with two Ig domains, the Ig3 and Ig2 domains, and a larger surface of the growth factor might be required to finely tune the receptor binding affinity. While the work presented here addresses many of the questions concerning structural features of the FGF protein family, future work is still needed to better understand interactions between FGF and the FGF receptor.

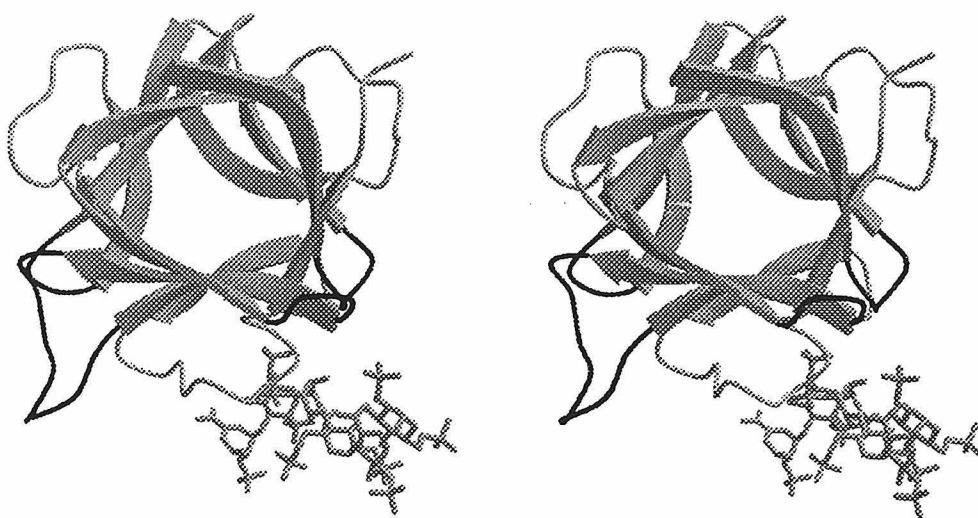


Fig. 3.5. Stereodigram of KGF with regions I, and II in black, and showing the position of heparin hexamer after superpositioning with the bFGF-heparin structure.

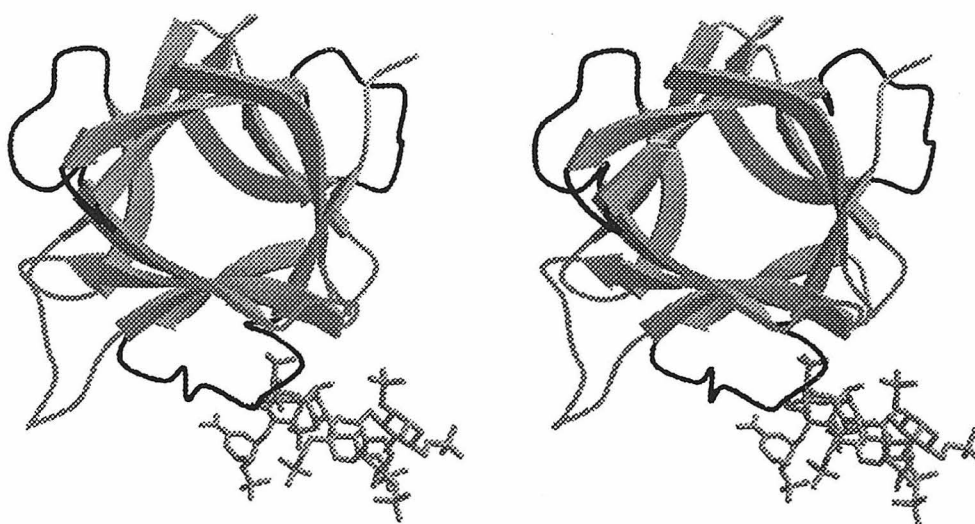


Fig. 3.6. Stereodescription of KGF showing in black loops that are likely to have structural variations due to size differences among members of the FGF family. The position of heparin after the superposition with the bFGF-heparin structure is shown.

References

- (1) W. Kabsch, *J. Appl. Cryst.* **21**, 916-924 (1988).
- (2) X. Zhu, et al., *Science* **251**, 90-93 (1991).
- (3) J. Navaza, *Methods Enzymol.* **276**, 581-594 (1997).
- (4) A. T. Bruger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
- (5) S. Bailey, *Acta Cryst D* **50**, 760-763 (1994).
- (6) T. A. Jones, *Methods Enzymol* **115**, 157 (1985).
- (7) W. Kabsch, *Acta crystallogr.* **A32**, 922-923 (1976).
- (8) P. J. Kraulis, *J. Appl. Cryst.* **24**, 946-950 (1991).
- (9) S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt, D.C. Rees, *Science* **271**, 1116-1120 (1996).
- (10) R. Reich-Slotky, D. Bonnef-Barkey, E. Shaoul, B. Bluma, C. M. Svahn, D. Ron, *J. Biol. Chem.* **269**, (1994).
- (11) N. J. Skelton, F. Aspiras, J. Ogez, T. J. Schall, *Biochemistry* **34**, (1995).

Appendix

**Role of the Active-Site Cysteine of *Pseudomonas aeruginosa* Azurin.
Crystal Structure Analysis of the Cu^{II}(Cys112Asp) Protein**

(Adapted from Faham et al. , JBIC in press.)

Introduction

Azurin is a relatively small (~15 kD) blue copper protein that shuttles electrons in the respiratory chains of certain denitrifying bacteria [1,2]. The copper in the wild-type (wt) protein is coordinated by two histidines (His46, His117) and a cysteine (Cys112) in a trigonal planar CuN_2S coordination geometry; in addition, there are axially disposed sulfur (Met121) and oxygen (Gly45 carbonyl) donor atoms that appear to interact weakly with the metal center [3-6]. Because of a relatively rigid active-site structure [7], the inner-sphere reorganization associated with azurin electron-transfer reactions is minimal [8-13].

Owing to the availability of *Pseudomonas aeruginosa* azurin from recombinant sources [14-16], site-directed mutagenesis has proven to be a popular and powerful technique for assessing directly the role of a residue or set of residues in tuning the properties of the protein [17]. The spectroscopic properties of a mutant in which the active-site cysteine had been replaced by an aspartic acid residue has been previously reported [18]. Although the Cys112Asp protein retained its ability to bind Cu^{II} strongly, its spectroscopic properties indicated the presence of a normal (type 2) copper at the active center. Consistent with electronic spectroscopic data, computer modeling based on results obtained from an NMR investigation of $\text{Co}^{\text{II}}(\text{Cys112Asp})$ azurin predicted a distorted square-pyramidal coordination geometry built from His46, His117, a bidentate Asp112, and a Gly45 backbone carbonyl group [19].

We report here that $\text{Cu}^{\text{II}}(\text{Cys112Asp})$ azurin reacts readily with certain mild reducing agents and that the resulting $\text{Cu}^{\text{I}}(\text{Cys112Asp})$ azurin is stable to metal loss; that titration of reduced *P. aeruginosa* cytochrome *c*₅₅₁ with oxidized Cys112Asp azurin indicates that the $\text{Cu}^{\text{II/I}}$ reduction potential is 180 mV vs. NHE; and that an analysis of the

crystal structure of the Cu^{II} mutant at 2.4 Å resolution confirms the predicted square-pyramidal coordination geometry.

Materials and methods

General

House distilled water was purified using a Barnstead NANOpure system. Protein samples were concentrated using Centricon concentrator units (Amicon) at 4 °C. Small-molecule contaminants in concentrated protein samples were removed by gel filtration through a PD-10 column (Pharmacia). Electronic absorption spectra were recorded on a modified Cary 14 spectrophotometer.

Protein Preparation

P. aeruginosa Cys112Asp azurin was prepared as described previously by J. Mizoguchi [18-20]. Cu^{II}(Cys112Asp) azurin was reconstituted from pure apo Cys112Asp azurin with the addition of a slight excess of aqueous CuSO₄. The Cu^{II}-reconstituted protein was treated with EDTA subsequent to FPLC purification in 10 mM DEA·HCl or EA·HCl (pH 8.8 or 9.0, respectively) on a Mono Q column (Pharmacia) using a linear NaCl gradient. Failure to sequester the excess Cu^{II} in the protein solution can result in FPLC chromatograms with multiple, poorly resolved azurin-containing fractions that still include varying amounts of excess Cu^{II}. Cu^{II}(Cys112Asp) azurin concentrations were determined using $\epsilon_{310} = 2.0 \times 10^3$ or $\epsilon_{754} = 1.0 \times 10^2$ M⁻¹cm⁻¹ [20]. *P. aeruginosa* cytochrome *c*₅₅₁ was obtained from Sigma. The Fe^{III} cytochrome was purified by FPLC in 10 mM DEA·HCl (pH 8.8) as described above for azurin [21].

Protein Crystallization and Data Collection

Cu^{II}(Cys112Asp) azurin crystals were grown by the hanging drop vapor diffusion method. 3.0 μ L protein (20 mg/mL in 0.1 M sodium acetate (pH 5.6)) were mixed with 3.0 μ L reservoir solution (0.1 M Tris-HCl (pH 8.0), 25% PEG-4000, 20 mM CuCl₂, 0.1 M LiNO₃). Colorless crystals grew over a few days at room temperature, reaching a final size of approximately 0.3 x 0.1 x 0.1 mm. The space group was determined to be *C*2 with unit cell constants $a = 56.5$, $b = 48.9$, $c = 95.5$ Å, $\beta = 94.0^\circ$, with two molecules in the asymmetric unit. Diffraction data were collected on a Siemens multiwire area detector at room temperature using X-rays generated by a rotating copper anode with Cu K α radiation ($\lambda = 1.54$ Å). The data were processed using the program XENGEN [22]. The data set used for the structure determination contained 8392 unique reflections to 2.4 Å (83% complete) merged from a total of 19753 observations with a merging *R* of 0.078. The highest shell (2.42 - 2.54 Å) was 51% complete, with a merging *R* of 0.17.

Structure Solution and Refinement

The structure was solved by molecular replacement techniques using the crystal structure of apoazurin [23] as a starting model. Using the X-PLOR package [24] for the molecular replacement calculations, two molecules were located in the asymmetric unit. These molecules are related by a non-crystallographic twofold axis along the crystallographic *z* axis, consistent with the results of self-rotation function calculations. The model was rebuilt with the TOM/FRODO [25] graphics program, and refined with X-PLOR. Clear electron density for the Asp112 was found in difference Fourier maps (Fig. A.1). Non-crystallographic symmetry restraints were used at intermediate stages of the refinement. Ultimately, these restraints were removed to allow the different monomers to refine

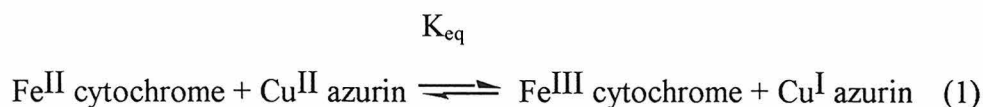
independently. The structure refined to $R = 0.20$ for data between 8.0 and 2.4 Å with $F > \sigma(F)$ with rms deviation from ideal values for bond distances and angles of 0.016 Å and 2.66°, respectively. The current Cys112Asp azurin model includes 1968 atoms assembled from 1960 protein atoms, 4 copper atoms, and 4 water molecules [26].

Results and Discussion

Cu^{I} (Cys112Asp) Azurin

Upon incubation of Cu^{II} (Cys112Asp) azurin with excess $[\text{Ru}^{\text{II}}(\text{NH}_3)_6]^{2+}$, followed by gel filtration, a colorless azurin-containing solution was isolated [27]. The absorption spectrum of this solution is markedly different from that of the starting Cu^{II} (Cys112Asp) azurin [20]. Notably, both the imidazole-to- Cu^{II} charge-transfer ($\lambda_{\text{max}} = \sim 310$ nm) and the Cu^{II} ligand-field ($\lambda_{\text{max}} = 754$ nm) bands are absent in the $[\text{Ru}^{\text{II}}(\text{NH}_3)_6]^{2+}$ -treated sample, thereby indicating reduction of Cu^{II} to Cu^{I} . These Cu^{II} -dependent absorption bands reappear slowly over several hours upon standing in air. Importantly, the absorption spectrum of Cu^{I} (Cys112Asp) azurin is remarkably similar to that of the corresponding Cu^{I} protein [20]; both reduced proteins show increased near-UV absorptions that can be ascribed to Cu^{I} -centered electronic transitions [28]. In light of our finding that the copper ion is not removed by gel filtration of the $[\text{Ru}^{\text{II}}(\text{NH}_3)_6]^{2+}$ -treated Cys112Asp azurin and that the absorption spectrum of the resulting sample is very similar to that of the reduced WT protein, it is likely that Cu^{I} remains coordinated in the Cys112Asp active site.

Addition of oxidized Cys112Asp azurin to a solution of reduced *P. aeruginosa* cytochrome *c*₅₅₁ is accompanied by changes in the Q-band region of the heme absorption spectrum that are attributable to the following equilibrium reaction:



Analysis of a titration experiment in terms of eq 1 gave $E^\circ(\text{Cu}^{\text{III/I}}) = 180 \pm 10$ mV vs. NHE (pH 7.0) [20]. Although the reduction potential of $\text{Cu}^{\text{II}}(\text{Cys112Asp})$ mutant is significantly lower than that of WT azurin (310 mV vs. NHE at pH 7.0) [29], it still falls within the range reported for blue proteins; in fact, the potential is the same as that of stellacyanin [30]. A decrease in reduction potential observed upon substitution of Cys112 with Asp would be expected based on the hardness of the carboxylate oxygen atoms and the change to a Cu^{II} -stabilizing tetragonal coordination environment. However, our finding that the potential remains relatively high suggests that the hydrophobic nature of the azurin active site [31] plays a key role in tuning the potentials of WT and mutant azurins.

Numerous studies have shown that the His117 ligand of azurin is the physiological port of entry and exit of electrons [32-34], in effect bypassing the highly coupled bridge to the copper center provided by the Cys112 thiolate [5,10]. Our work on the Cys112Asp mutant confirms the key role that His117 plays in the ET pathway between azurin and cytochrome *c*₅₅₁: even in the absence of the cysteine thiolate, relatively facile electron transfer takes place between these two proteins.

Overall Structure

The crystal structure of $\text{Cu}^{\text{II}}(\text{Cys112Asp})$ azurin reveals the presence of two crystallographically independent molecules in the asymmetric unit (Fig. A.2). These two azurin molecules are nearly identical, with a 0.3 Å rms deviation between their main-chain atom positions. The interface between the two azurin molecules in the asymmetric unit is lined with residues (Met13, Leu39, Pro40, Val43, Met44, Phe114, Pro115, Ala119, and Leu120) that define the hydrophobic patch of *P. aeruginosa* azurin (Fig. A.2) [31]. The distance between the two active site coppers in the dimer is 16.8 Å. As in the WT protein [35,36], the His117 side chain, located in the middle of the hydrophobic patch, remains partially solvent exposed.

During the course of the structure determination, a strong peak was found near the N-terminus in ($F_o - F_c$) difference Fourier maps. Since CuCl_2 was used for crystal growth, this additional density was modeled as copper. The average B factor for the exogenous copper refined to 29.04 Å²; The average B factor for all atoms refined to 22.7 Å². The exogenous copper center is coordinated by the amino nitrogen and carbonyl oxygen atoms of Ala1, the imidazole N^ε of His83 from an adjacent azurin molecule, and a water molecule in a distorted square planar geometry (Fig. A.3). The presence of this site explains the role of copper in the crystallization of $\text{Cu}^{\text{II}}(\text{Cys112Asp})$ azurin, since it acts as a crosslinking agent between neighboring protein molecules (Fig. A.3) [37]. The Cys112Asp azurin dimer structure is distinct from all other azurin structures [35,36,38]; this difference may result partly from the orientational requirements imposed by the surface-bound copper ion.

The crystal structures of the low (5.5) and high (9.0) pH forms of WT azurin show that the Pro36-Gly37 peptide bond undergoes significant structural rearrangement

between the two pH values [35]. A hydrogen bond is present in the low pH structure between the Pro36-Gly37 peptide carbonyl group and a protonated N^ε of His35. At high pH, this hydrogen bond is broken as the carbonyl group flips out, and the His35 side chain becomes deprotonated. Although the orientation of the Pro36-Gly37 peptide carbonyl group in the Cys112Asp azurin structure is intermediate to those of the low and high pH forms of the WT protein, a hydrogen bond between the carbonyl group and His35 in the mutant structure does not exist. Therefore, we have used the Cu^{II} WT azurin structure at pH 9.0 (PDB code 5azu) to make quantitative comparisons with that of Cu^{II}(Cys112Asp) azurin.

The main-chain atoms of the WT and Cys112Asp azurin monomers are nearly superimposable, except at the N-terminus. Omitting residues Ala1 and Glu2, the rms deviation in the main-chain atom positions between the mutant and WT structures is 0.5 Å. A localized disparity in this region is not surprising, as the flexibility is generally greater at the termini of proteins. Furthermore, as described above, an exogenous Cu^{II} ion is coordinated both to the carbonyl oxygen and the terminal amino group of Ala1. There is a rotation of ~180 degrees about the C^α-N bond of Cys3 in Cys112Asp azurin. As a result, the terminal amino group, in addition to coordinating the copper ion, is hydrogen bonded to O^γ(Ser25) (Fig. A.3). These findings suggest that the amino terminus of azurin is conformationally flexible. It is noteworthy that the third residue in the azurin sequence is a cysteine that is involved in a disulfide bond with Cys26. Therefore, the Cys3-Cys26 disulfide bond may play a critical role in stabilizing the native β-barrel structure of the protein, which is somewhat larger than those of other cupredoxins. In fact, mutagenesis of these cysteine residues in *P. aeruginosa* azurin has demonstrated the importance of the disulfide in preserving the native polypeptide fold [39].

Active Site

The active-site copper of Cu^{II}(Cys112Asp) azurin is coordinated to the N^δ atoms of His46 and His117, the backbone carbonyl oxygen of Gly45, and both carboxylate oxygens of Asp112 in a distorted square-pyramidal geometry (Table A.1). The bidentate Asp112 carboxylate-copper interaction is distinguished by its asymmetric coordination such that one of the metal-oxygen bonds is ~0.6 Å longer than the other, and the displacement of the metal by ~1 Å out of the plane defined by the carboxylate group of Asp112. The ~3.5 Å interatomic distance precludes the existence of a bond between the thioether sulfur of Met121 and the copper center. Differences in coordination geometry from the active-site copper in the WT protein are achieved by movements of Met121 away from the copper, and of the copper and Gly45 toward each other (Fig. A.4).

The distorted square-pyramidal coordination geometry of Cu^{II}(Cys112Asp) azurin that was proposed from spectroscopic and computer modeling studies [19] is confirmed in the crystal structure. The structure features a metal center with a bidentate, asymmetric, out-of-plane interaction with the Asp112 side-chain carboxylate group and an axial ligand provided by the backbone carbonyl group of Gly45. It appears that the polypeptide fold of azurin is sufficiently rigid to force the metal-bidentate carboxylate unit to adopt a nonplanar structure [40]. Therefore, this unusual carboxylate coordination is an example of a rack-induced distortion [41]. The two WT Cys112 side chain-to-main chain hydrogen bonds are preserved in the mutant active site. However, instead of two hydrogen bonds to a single sulfur atom (S^γ(Cys112)–NH(Asn47) and S^γ(Cys112)–NH(Phe114)), the bonds are split, one to each of the two Asp112 carboxylate oxygens (O^{δ1}(Asp112)–NH(Asn47) and O^{δ2}(Asp112)–NH(Phe114)).

The introduction of two extra atoms (the carboxylate oxygens) from the Asp112 side chain produces many small perturbations in the active-site structure. Notably, the entire His117 side chain has translated toward the active-site cavity to a position closer to the Gly45 carbonyl group, thereby withdrawing the solvent-exposed portion of the His117 imidazole ring slightly away from the bulk solvent. The relatively long Met121 side chain has moved so that the thioether sulfur atom is located farther from the active-site core. Finally, the entire Asp112 side chain is pulled away from the active-site cavity, perhaps to lessen the severity of the out-of-plane binding of the copper ion, or merely to accommodate the longer side chain.

Table A.1. Active-Site Interatomic Distances for *P. aeruginosa* Cu^{II}(Cys112Asp) and WT Azurins.

atomic pair	distance (Å)	distance (Å)
	(Cu ^{II} (Cys112Asp) azurin) ^a	(Cu ^{II} WT azurin) ^b
Cu-S ^γ (Cys112)	-	2.3
Cu-O ^{δ1} (Asp112)	2.5	-
Cu-O ^{δ2} (Asp112)	1.9	-
Cu-N ^δ (His46)	2.1	2.1
Cu-N ^δ (His117)	2.0	2.0
Cu-O(Gly45)	2.6	3.0
Cu-S ^δ (Met121)	3.5	3.1

^a Average for the two crystallographically independent Cys112Asp azurin molecules. ^b

Average for the four crystallographically independent WT azurin molecules at pH 9.0 (PDB code 5azu) [35].

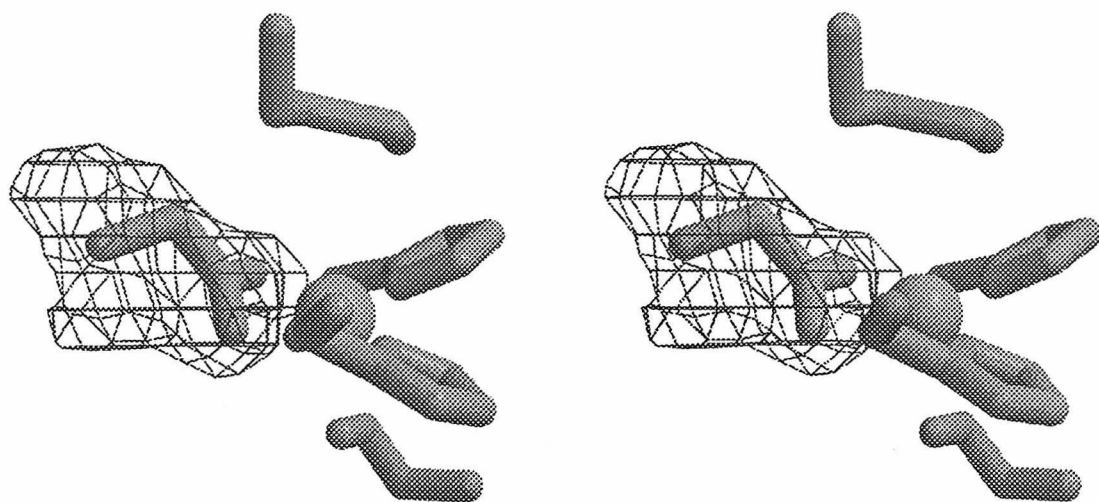


Figure A.1. Stereodescription of the active site, showing a Fourier difference map ($F_o - F_c$) at 3.0 overall rms value, where residue 112 was deleted from the model.

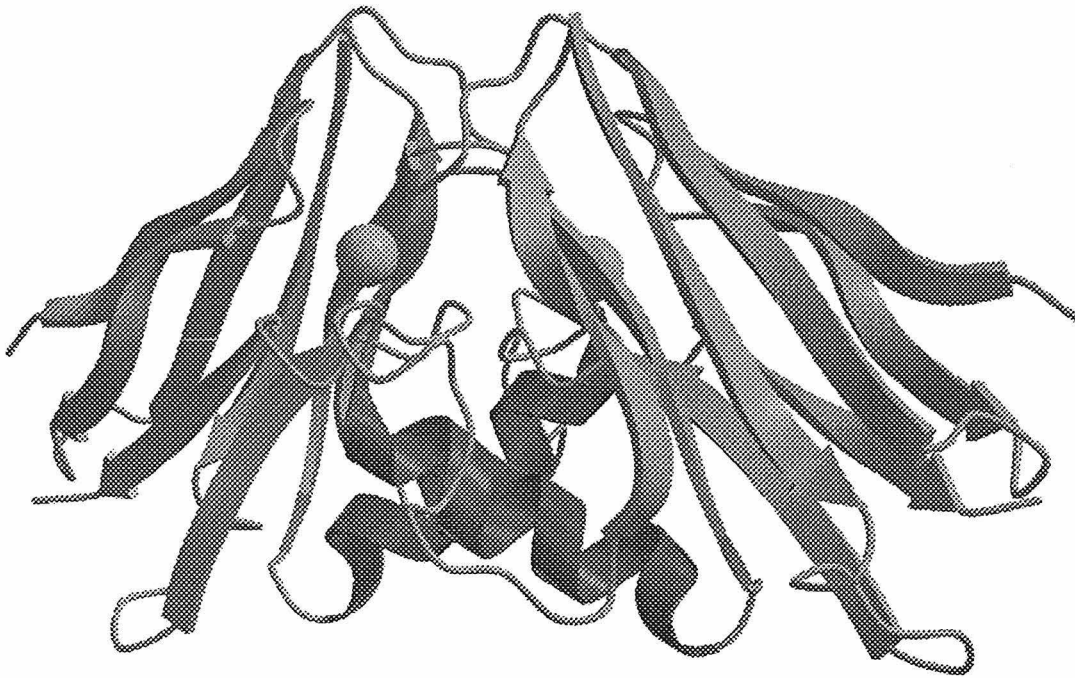


Figure A.2. Ribbons diagram of the Cu^{II}(Cys112Asp) azurin dimer that constitutes an asymmetric unit.

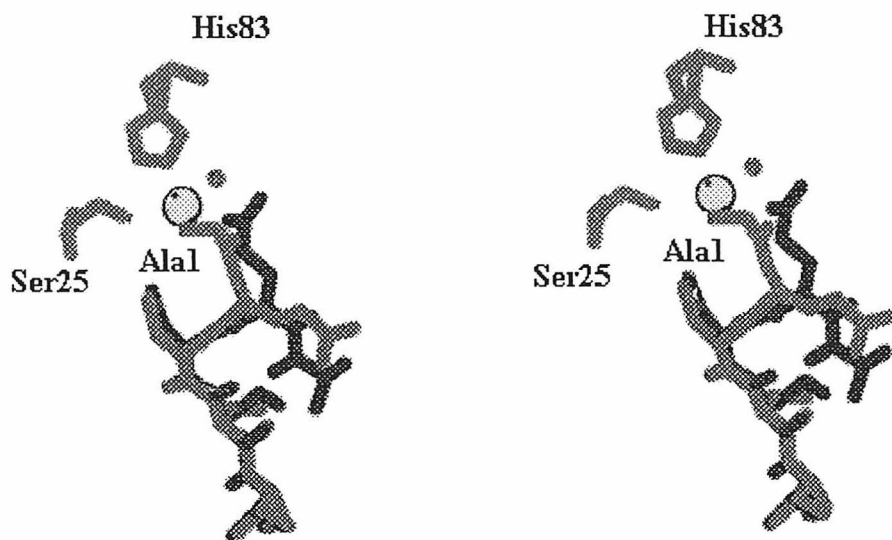


Figure A.3. Stereodigram of the amino termini of native (black) and Cys112Asp azurin (gray), showing the different conformations adopted by the first two residues, the hydrogen bonding between the amino nitrogen and the hydroxyl group of Ser25, and the coordination geometry of the exogenous copper.

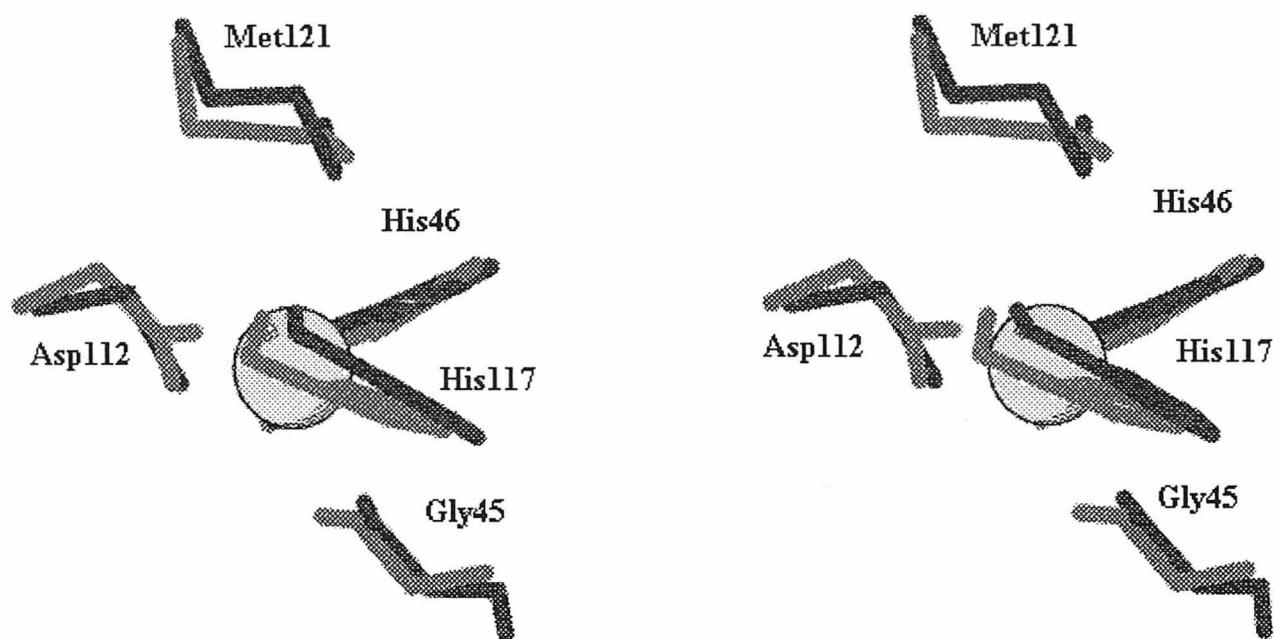


Figure A.4. Stereodigram of the coordination structure of the active-site copper of Cu^{II}(Cys112Asp) azurin (gray) superimposed on the active site of the native azurin structure (black). See Table A.1 for a list of interatomic distances.

References

- (1) Sykes AG (1991) *Adv. Inorg. Chem.* 36:377-408.
- (2) Adman ET (1991) *Adv. Protein Chem.* 42:145-197.
- (3) Karlsson BG, Aasa R, Malmström BG, Lundberg LG (1989) *FEBS Lett.* 253:99-102.
- (4) Di Bilio AJ, Chang TK, Malmström BG, Gray HB, Karlsson BG, Nordling M, Pascher T, Lundberg LG (1992) *Inorg. Chim. Acta* 198-200:145-148.
- (5) Solomon EI, Penfield KW, Gewirth AA, Lowery MD, Shadle SE, Guckert JA, Lacroix LB (1996) *Inorg. Chim. Acta* 243:67-78.
- (6) Guckert JA, Lowery MD, Solomon EI (1995) *J. Am. Chem. Soc.* 117:2817-2844.
- (7) Shepard WEB, Anderson BF, Lewandoski DA, Norris GE, Baker EN (1990) *J. Am. Chem. Soc.* 112:7817-7819.
- (8) Gray HB, Winkler JR (1996) *Annu. Rev. Biochem.* 65:537-561.
- (9) Winkler JR, Gray HB (1992) *Chem. Rev.* 92:369-379.
- (10) Regan JJ, Di Bilio AJ, Langen R, Skov LK, Winkler JR, Gray HB, Onuchic JN (1995) *Chem. Biol.* 2:489-496.
- (11) Langen R, Chang I-J, Germanas JP, Richards JH, Winkler JR, Gray HB (1995) *Science* 268:1733-1735.
- (12) Connick WB, Di Bilio AJ, Hill MG, Winkler JR, Gray HB (1995) *Inorg. Chim. Acta* 240:169-173.
- (13) Langen R, Colon JL, Casimiro DR, Karpishin TB, Winkler JR, Gray HB (1996) *JBIC* 1:221-225.
- (14) Karlsson BG, Pascher T, Nordling M, Arvidsson RHA, Lundberg LG (1989) *FEBS Lett.* 246:211-217.

- (15) van de Kamp M, Hali FC, Rosato N, Finazzi Agro A, Canters GW (1990) Biochim. Biophys. Acta 1019:283-292.
- (16) Chang TK, Iverson SA, Rodrigues CG, Kiser CN, Lew AYC, Germanas JP, Richards JH (1991) Proc. Natl. Acad. Sci. USA 88:1325-1329.
- (17) Canters GW, Gilardi G (1993) FEBS Lett. 325:39-48.
- (18) Mizoguchi TJ, Di Bilio AJ, Gray HB, Richards JH (1992) J. Am. Chem. Soc. 114:10076-10078.
- (19) Piccioli M, Luchinat C, Mizoguchi TJ, Ramirez BE, Gray HB, Richards JH (1995) Inorg. Chem. 34:737-742.
- (20) Mizoguchi TJ (1996) Ph.D. Thesis, California Institute of Technology.
- (21) Purification of commercial cytochrome *c*₅₅₁ revealed two major peaks in the FPLC. Samples from both fractions exhibited superposable absorption spectra indicative of Fe^{III} cytochromes *c*. Protein from the larger peak, which eluted at higher salt concentrations, was used in the current work owing to its greater availability.
- (22) Howard AJ, Gilliland GL, Finzel BC, Poulos TL, Ohlendorf DH, Salemme FR (1987) J. Appl. Cryst. 20:383-387.
- (23) Adman ET Unpublished results: apoazurin from *P. aeruginosa* was provided by H. Engeseth and D.R. McMillin; the model was refined using 16640 data between 16.0-2.5 Å collected on a Picker diffractometer, CuKα radiation. Space group P2₁2₁2₁, with cell dimensions 57.26x81.11x110.61 Å. Refined with X-PLOR (parameter file PARAM19) R=0.180; for 4 molecules in asymmetric unit 212 solvent total (~50 per monomer).
- (24) Brünger AT, Kuriyan J, Karplus M (1987) Science 235:458-460.
- (25) Jones TA (1985) Methods Enzymol. 115:157-171.

- (26) The crystallographic coordinates of Cu^{II}(Cys112Asp) azurin have been deposited in the Brookhaven National Laboratory Protein Data Bank.
- (27) Similar treatment using Cu^{II} WT azurin results in the isolation of the colorless Cu^I WT protein.
- (28) Tamilarasan R, McMillin DR (1986) *Inorg. Chem.* 25:2037-2040.
- (29) Pascher T, Karlsson BG, Nordling M, Malmström BG, Vänngård T (1993) *Eur. J. Biochem.* 212:289-296.
- (30) Ikeda O, Sakurai T (1994) *Eur. J. Biochem.* 219:813-819.
- (31) Tsai L-C, Sjölin L, Langer V, Pascher T, Nar H (1995) *Acta Crystallogr., Sect. D* 51:168-176.
- (32) Van de Kamp M, Silvestrini MC, Brunori M, Van Beeumen J, Hali FC, Canters GW (1990) *Eur. J. Chem.* 194:109-118.
- (33) Van Pouderoyen G, Mazumdar S, Hunt NI, Hill HAO, Canters GW (1994) *Eur. J. Biochem.* 222:583-588.
- (34) Gorren ACF, den Blaauwen T, Canters GW, Hopper DJ, Duine JA (1996) *FEBS Lett.* 381:140-142.
- (35) Nar H, Messerschmidt A, Huber R, Van de Kamp M, Canters GW (1991) *J. Mol. Biol.* 221:765-772.
- (36) Adman ET, Jensen LH (1981) *Israel J. Chem.* 21:8-12.
- (37) Copper crosslinking near the N-termini of protein molecules has been observed in crystals of thioredoxin: Holmgren A, Soderberg B, Ekland H, Branden C-I (1975) *Proc. Natl. Acad. Sci. USA* 72:2305-2309.
- (38) Day MW (1995) Ph.D. Thesis, California Institute of Technology.

- (39) Bonander N, Karlsson BG, Vänngård T (1995) *Biochim. Biophys. Acta* 1251:48-54.
- (40) For a discussion of metal binding to isolated carboxylate groups in proteins, see:
Glusker JP, (1991) *Adv. Protein Chem.* 42:1-76.
- (41) Malmström BG (1994) *Eur. J. Biochem.* 223:711-718.