Complex with a Phage Display-Derived Peptide Provides Insight into the Function of Insulin-like Growth Factor I‡

Michelle L. Schaffer, Kurt Deshayes, Gerald Nakamura, Sachdev Sidhu, and Nicholas J. Skelton*

Department of Protein Engineering, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080

Received March 7, 2003; Revised Manuscript Received June 12, 2003

ABSTRACT: The dramatic improvement in the NMR spectra of insulin-like growth factor I (IGF-I) in the presence of a peptide identified from a phage display library has allowed for the first time the determination of a high-resolution solution structure for much of IGF-I. The three helices of IGF-I in this complex have an arrangement similar to that seen in high-resolution crystal structures of IGF-I and insulin, although there are differences in the conformation and precise location of helix 3. A cluster of hydrophobic and basic side chains within the turn−helix motif of the peptide contact a hydrophobic patch on helices 1 and 3 of IGF-I. The importance of this patch for tight binding was verified using alanine scanning mutagenesis of the peptide in two different phage display formats. Consistent with its antagonistic activity, the peptide binds to a region implicated by mutagenesis studies to be important for association with IGF binding proteins (IGFBPs). The ability of the peptide to also inhibit signaling has important implications for the manner in which IGF-I interacts with its receptor. Interestingly, the peptide uses the same binding site as detergent and a fragment of IGFBP-5 identified in other IGF-I complexes. The ligand-induced structural variability of helix 3 in these complexes suggests that exchange between such conformations may be the source of the dynamic nature of free IGF-I and likely has functional significance for the ability of IGF-I to recognize two signaling receptors and six binding proteins with high affinity.

Insulin and the insulin-like growth factors (IGF-I[1] and IGF-II) belong to the insulin superfamily of growth factors that control glucose metabolism, cell growth, and embryonic development. Although multiple biological activities have been attributed to these proteins, insulin is primarily involved in metabolic control whereas the activity of IGF-I is more relevant in growth regulation and development (*1*). Insulin, IGF-I, and IGF-II are all produced as single polypeptide chains containing three intrachain disulfide bonds. In the case of proinsulin, a central region is proteolytically excised to produce the mature protein that contains separate A- and B-chains covalently attached by two of the disulfide bonds. Thus, IGF-I has a C-region loop and a C-terminal D-region tail that have no counterpart in insulin. The so-called A- and B-regions of IGF-I share ∼50% sequence identity with insulin (Figure 1).

Insulin and IGF are usually resident in the serum and exert their biological activity by initiating a complex intracellular signal cascade on association with cell surface $\alpha_2\beta_2$ heterotetrameric receptors (*2*). Although detailed structural information of the ligand−receptor interaction is not available, biophysical characterization of the association between many insulin and insulin receptor (I-R) variants has led to a detailed multistep model for the ligand association and activation of I-R (*1*, *3*). Although not studied in such detail, the interaction of IGF-I with the IGF-I receptor (IGF-R) appears to proceed by a similar mechanism (*4*). IGF-I is further regulated by its association with one of six serum IGF binding proteins [IGFBP-1−6 (*5*)] that sequester IGF in a high-affinity complex, in some cases involving a third protein (the acid-labile subunit or ALS) (*6*). This process protects IGF from rapid degradation or elimination from the serum and keeps the circulating concentrations of biologically active, uncomplexed IGF low (*7*, *8*). Moreover, the ternary complex provides a reservoir of available IGF-I (from which IGF-I can be liberated by binding protein proteolysis) and may also be involved in transport of the growth factor to particular vascular locations (*8*).

Numerous crystallographic and solution-state structural studies have been performed on insulin over the last 3 decades (e.g., see refs *9*−*12*). These studies show that the A-chain contains two short helices that pack onto a longer, third helix in the B-chain. The insulin monomers (A- and B-chain) associate in the crystalline lattice to form dimers and Zn^{2+}-mediated hexamers, forms thought to be of biological importance for storage in the pancreas but which must disassociate into monomers for activity (*11*). Although defined at high resolution, the conformation observed for the monomer is not thought to be functionally relevant and "opening" of the B-chain C-terminus to reveal the termini of the A-chain is required for receptor binding (*10*, *13*).

‡ Atomic coordinates have been deposited in the Protein Data Bank (Accession Code 1PMX).

* To whom correspondence should be addressed [tel, (650) 225-6402; fax, (650) 225-3734; e-mail, skelly@gene.com].

[1] Abbreviations: 2QF-COSY, double-quantum-filtered spectroscopy; ELISA, enzyme-linked immunosorbent assay; Fmoc, 9-fluorenylmethoxycarbonyl; HPLC, high-performance liquid chromatography; HSQC, heteronuclear single-quantum coherence; IGF-I, insulin-like growth factor I; IGFBP, IGF binding protein; IGF-R, type 1 insulin-like growth factor I receptor; I-R, insulin receptor; NMR, nuclear magnetic resonance spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; RMS, root mean squared; ROESY, rotating frame Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy.

10.1021/bi034386c CCC: $25.00
Published on Web 07/17/2003

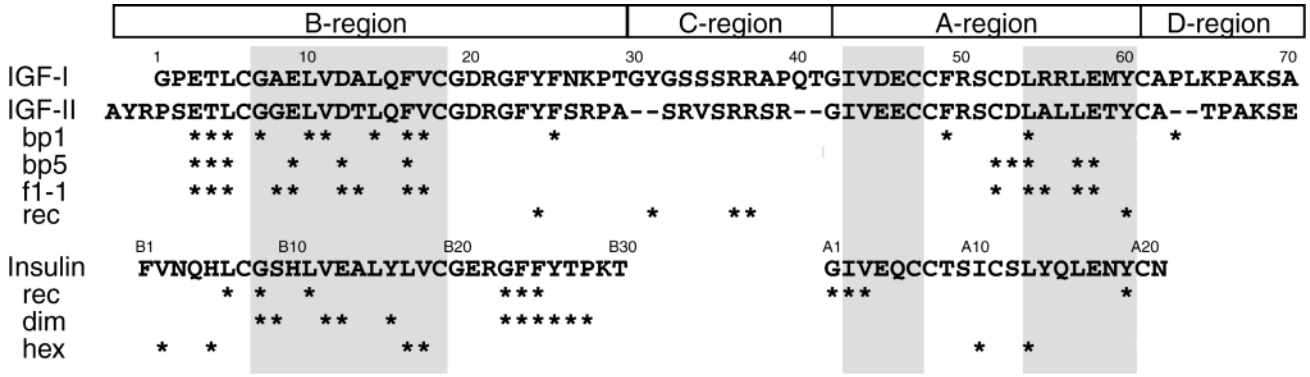

FIGURE 1: Sequence alignment of IGF-I, IGF-II, and insulin. The three helices are indicated by the shaded boxes. Residues of IGF-I that have been implicated in binding to IGFBP-1 (bp1), IGFBP-5 (bp5), IGF-F1-1 (f1-1), and IGF-R (rec) by structural or mutagenesis studies are indicated by the stars below the sequence. Residues of insulin that are involved in the dimer and hexamer interfaces of the Zn^{2+} form of insulin or are implicated in I-R binding (rec) are similarly indicated below the sequence.

Structural studies on IGF-I have been less productive due to its dynamic nature in solution (*14*, *15*). Structures determined in solution at low pH for several variants of IGF-I identified the three A- and B-region helices, but the remainder of the protein and the relative orientation of the helices were not well-defined (*14*−*17*). A higher precision structure has been reported for IGF-II at low pH in solution (*18*). More recently, several high-resolution structures have been reported for IGF-I in complex with detergent molecules or a fragment of IGFBP-5 (*19*−*21*). Much of the A- and B-regions in these structures is arranged in a fashion very similar to the so-called T form of insulin, although there are some differences that have been ascribed functional significance (*20*).

To learn more about this important signaling pathway, we have used phage display to obtain specific peptide ligands to its component proteins. Initial efforts focused on identifying peptides that bind to IGFBP-1 and block IGF-I binding (*22*) and characterizing the molecular requirements of this interaction (*23*). Molecules of this type hold potential as "IGF-I agonists" in the treatment of metabolic disease, since they prevent association with binding protein and facilitate improved interaction of IGF-I with the cell surface signaling receptors (*22*, *24*). More recently, we have used an optimized phage selection approach to identify peptides that bind to IGF-I (*25*). One such peptide, designated IGF-F1-1, binds to IGF-I with low micromolar affinity, has a well-defined structure by itself in solution, and is able to block both binding protein association and receptor signaling (*25*).

In the present work, we have characterized in more detail the interaction between IGF-F1-1 and IGF-I. Two independent phage display-based methods have been used to identify a hydrophobic and basic patch on the surface of the peptide that is required for binding to IGF-I. Moreover, addition of peptide leads to a dramatic improvement in the quality of the NMR spectra of IGF-I, thereby allowing determination of the structure of the complex by triple resonance NMR methods at high resolution. The peptide is found to bind at a complementary hydrophobic and acidic patch between helix I and helix 3 of IGF-I, a location that readily explains the ability of the peptide to block association with IGFBPs and suggests possibilities for its ability to reduce IGF-I receptor binding. Comparison to the IGF-I structures in complex with detergent and mini-IGFBP-5 provides further insights into the function and behavior of IGF-I.

MATERIALS AND METHODS

Preparation of Labeled IGF-I. Esherichia coli 43E7 cells were transformed with the IGF-I-expressing vector pBKIGF2B (*26*). Labeled protein was produced as described previously (*27*). Briefly, overnight starter cultures were added to rich media containing either CELTONE-N or CELTONE-CN (Martek) for uniform ^{15}N or $^{15}N/^{13}C$ labeling, respectively. In addition to the CELTONE powder, the media consisted of 120 mM triethanolamine (pH 7.4), 1.6 mM $MgSO_4$, 20 mM $^{15}NH_4Cl$, 50 mM KCl, 20 mM NaCl, 50 μg/mL carbenecillin, and 0.3% glucose (3 g/L; ^{13}C in the case of the $^{15}N/^{13}C$ media). After 18 h at 37 °C the cultures reached an A_{600} of 3.3, and the cells were pelleted and washed with a solution of 120 mM triethanolamine (pH 7.4), 1.6 mM $MgSO_4$ 50 mM KCl, 20 mM NaCl, and 50 μg/mL carbenecillin. The cells were resuspended in minimal media (as above without CELTONE powder) to an A_{600} of 3.0 and grown at 37 °C for 24 h. The cells were pelleted, resuspended in lysis buffer (25 mM Tris and 5 mM EDTA, pH 7.5), and passed through a microfluidizer three times. The refractile bodies were washed with lysis buffer and lysis buffer containing 1% laurylsarcosine. The pellet was then resuspended in refolding buffer (50 mM CAPS, 2 M urea, 100 mM NaCl, 20% methanol, pH 10.2) to which was added 20 mM DTT. The suspension was stirred at room temperature for 2 h, diluted 10-fold with refolding buffer, and stirred overnight at 4 °C. After addition of 5% acetic acid (1 mL) the insoluble debris was removed by centrifugation. The correctly folded portion of IGF-I was then purified by reversed-phase high-performance liquid chromatography (HPLC) on a C18 column as described (*28*). The protein purity and integrity were verified by analytical HPLC, SDS−PAGE, and mass spectrometry. The overall yield was approximately 10 mg of IGF-I/L of media.

Peptide Synthesis. Peptides were synthesized using standard 9-fluorenylmethoxycarbonyl (Fmoc) protocols, cleaved off the resin with 2.5% triisopropylsilane and 2.5% H_2O in trifluoroacetic acid (TFA), and purified by HPLC, as described previously (*25*). The purity and mass of each peptide were verified by liquid chromatography/mass spectrometry (LC/MS).

NMR Spectroscopy. NMR samples contained 1.4 mM (^{15}N-labeled) or 1.0 mM ($^{15}N/^{13}C$-labeled) IGF-I with 25 mM acetate-d_3 buffer (pH 5.1) in 93% ^1H_2O/7% ^2H_2O and 20

μM 1,4-dioxane as a chemical shift reference. NMR spectra were acquired at 40 °C on Bruker DRX600 and DRX800 spectrometers equipped with triple resonance, triple axis actively shielded gradient probes. The complex was formed by addition of aliquots of the phage-derived peptide IGF-F1-1 (RNCFESVAALRRCMYG-NH_2) until no further changes were seen in peak positions in $^1H-^{15}N$ HSQC spectra. Backbone resonance assignments for IGF-I were obtained from the following double and triple resonance experiments acquired at 600 MHz in H_2O solution, as described (*29*, *30*): $^1H-^{15}N$ 3D NOESY-HSQC; $^1H-^{15}N$ 3D TOCSY-HSQC, 3D HNCA, 3D HNCO, 3D CBCA(CO)-NH, and 3D HBHA(CBCACO)NH. A ^2H_2O sample was prepared by lyophilization of the $^{15}N/^{13}C$-labeled IGF-I complex and resuspension in 99.995% D_2O. A $^{15}N-^1H$ HSQC spectrum recorded 45 min after addition of ^2H_2O was used to identify amide protons undergoing slow exchange with solvent. IGF-I side chain assignments were completed from a 3D HCCH-TOCSY spectrum acquired at 800 MHz on this sample in ^2H_2O. Peptide resonance assignments were obtained from ^{14}N-filtered 2D ^1H NOESY and TOCSY spectra at 600 MHz in H_2O solution and ^{12}C-filtered ^1H NOESY and TOCSY spectra at 800 MHz in ^2H_2O solution. Details of these experiments are presented in supplementary Table 1 (see Supporting Information).

Determination of Structure. Distance restraints were obtained from analysis of the following NOESY spectra acquired at 800 MHz: 3D $^1H-^{15}N$ NOESY-HSQC, 2D ^{14}N-filtered ^1H NOESY, 3D ^{13}C-edited NOESY, and 2D ^{12}C-filtered ^1H NOESY. Intermolecular interactions were identified unambiguously in a 3D ω_1-^{12}C-filtered, ω_2-^{13}C-edited NOESY spectrum. Initial NOESY peak assignments were made on the basis of the assigned resonance positions and the structure of IGF-I in complex with the detergent deoxy big CHAPS (*19*), followed by several rounds of structure calculation and manual restraint checking and peak assignment. Dihedral angle restraints were obtained from analysis of 3D $^{15}N-^1H$ HNHA (ϕ), 3D $^{15}N-^1H$ HNHB, and 3D $^{15}N-^1H$ TOCSY-HSQC (χ_1) spectra. Additional conservative backbone dihedral angle restraints were obtained from analysis of backbone chemical shifts with the program TALOS (*31*). Restraints were applied for good fits to the chemical shifts (as defined by the program) with the allowed range being the TALOS-defined mean ± the larger of 30° or 3 times the TALOS-calculated standard deviation. Restraints for 12 hydrogen bonds were imposed within the helices (8, 2, and 2 in helices 1, 2, and 3, respectively) on the basis of slow amide proton exchange with solvent. Structures were calculated using the program CNX (v2000.1; Accelrys, San Diego). A total of 100 structures were calculated using torsion angle dynamics followed by Cartesian dynamics and minimization. The 20 structures of lowest restraint violation energy were chosen to represent the solution structure of the IGF-I in complex with IGF-F1-1. Details of the input restraints and structural statistics are presented in Table 1. Superposition of the ensemble was made using the backbone heavy atoms of residues 3−26 and 42−63 of IGF-I and 2′−14′ of IGF-F1-1. Comparison to other IGF-I structures was made by first superposing the backbone heavy atoms of residues helix 1, the B-region β-strand, and helix 2 (9−18, 23−26, and 42−49); these residues were judged most similar in all structures on the basis of the data in Figure 5 (see below).

Table 1: Summary of Input Restraints and Structural Statistics for the Final Ensemble of IGF-I in Complex with IGF-F1-1

parameter	ensemble
input restraints	
NOE total	905
intraresidue	146
sequential	203
medium range	232
long range	237
intermolecular	87
hydrogen bond (two per H-bond)	24
dihedral angles, total	139
ϕ	72
ψ	44
χ_1	23
violations	
RMSD from experimental restraints	
NOE distance (905)	0.0049 ± 0.0008
dihedral (139)	0.086 ± 0.027
NOE distance violations	
no. >0.01 Å	25.3 ± 6
no. >0.1 Å	0.25 ± 0.44
mean maximum violation (Å)	0.08 ± 0.03
dihedral angle violations	
no. >0.1 deg	11.0 ± 2.4
mean maximum (deg)	0.9 ± 0.3
RMSD from idealized geometry	
bonds (Å)	0.0021 ± 0.0001
angles (deg)	0.36 ± 0.01
impropers (deg)	0.15 ± 0.01
energy components (kcal·mol^{-1})	
total	62.0 ± 4.6
NOE	1.08 ± 0.32
CDIH	0.15 ± 0.10
bonds	5.9 ± 0.8
angles	48.1 ± 3.1
impropers	2.3 ± 0.3
van der Waals	4.5 ± 0.6
stereochemistry	
Ramachandran (%)	
favored	82.1
allowed	17.3
generous	0.6[a]
disallowed	0
structural precision	
mean RMSD to mean structure	
backbone IGF 3−26, 42−63 + F1-1 3−15	0.35 ± 0.06
heavy IGF 3−26, 42−63 + F1-1 3−15	0.89 ± 0.07
backbone secondary IGF 3−26, 42−63	0.27 ± 0.05
heavy secondary IGF 3−26, 42−63	0.75 ± 0.07
backbone vs X-ray 3−36, 42−63[b]	1.01

[a] Those residues falling in the generous region were in the ill-defined C- and D-regions of IGF-I; no more than three members of the ensemble had any one residue in this region. [b] Comparison to IGF-I in complex with deoxy big CHAPS (PDB accession 1imx).

Peptide Alanine Scan Analogues on Phage. An alanine scan was performed on the IGF-F1-1 peptide in a monovalent phage display format using the phagemid vector phGHamg3 (*32*) modified using Kunkel mutagenesis (*33*). Oligonucleotides were designed and synthesized to encode either wild-type IGF-F1-1 or single alanine mutations (glycine in the cases where alanine occurs in the wild-type sequence) for every residue except cysteine. The phagemid clones were verified by DNA sequencing. A competition ELISA was used to determine the relative IC_{50} for each alanine mutation compared to wild-type IGF-F1-1, as described previously (*25*). Briefly, IGF-I was coated on immunosorbent plates (Nunc Maxisorp) and washed with phosphate-buffered saline

(PBS) containing bovine serum albumin (BSA, 0.05%). Serial dilutions of IGF-I were incubated with the phage clones, followed by washing and incubation with horseradish peroxidase-conjugated anti-M13, and developed with a tetramethylbenzidine substrate. The assay was quenched with 1 M phosphoric acid, and the absorbance was read at 450 nm to determine the amount of bound phage remaining.

Shotgun Alanine Scan of Peptide. Libraries for combinatorial alanine scanning of IGF-F1-1 were constructed as described previously with a phagemid containing an IPTG-inducible promoter driving the expression of open reading frames containing an N-terminal STII secretion signal followed by IGF-F1-1, a linker (GGGSGGG), and finally the M13 gene-8 major coat protein (*34*, *35*). The degenerate codons described in Table 1 of Weiss et al. (*35*) were used to enable alanine scanning of the peptide. The selection protocol was identical to that used for initial selection of IGF-F1-1 (*25*). Briefly, the phage library was exposed to IGF-I coated on immunosorbent plates for 1 h and washed with phosphate-buffered saline (PBS) containing 0.05% Tween-20. The bound phage were eluted with glycine (pH 2.0), neutralized with 1.0 M Tris (pH 8.0), and then amplified in *E. coli* XL1-Blue cells. After two rounds of selection, a single-point phage ELISA was used to identify 192 clones positive for binding to immobilized IGF-I which were then sequenced (*25*). The program SGCOUNT was used to align the phage sequences and tabulate the occurrence of each amino acid type at each position within the peptide (*35*).

RESULTS

NMR Characterization of the Complex between IGF-F1-1 and IGF-I. IGF-I has been notoriously difficult to characterize by NMR spectroscopy due to a combination of self-association and internal dynamics that give rise to a range of line widths that are temperature, pH, and concentration dependent (*14*, *36*). Low pH and elevated temperature permitted assignment of most ^{1}H NMR resonances (*14*) and the determination of low-resolution structures (*14*, *15*). This poor behavior was confirmed in the present study where the ^{1}H−^{15}N HSQC spectrum of IGF-I at pH 5.0 and 40 °C contains 18 intense and 15 broad and weak peaks from backbone resonances: 64 backbone peaks are expected (supplementary Figure 1A). Moreover, the intense peaks occur over a very narrow range of H^N chemical shifts (8.2−8.6 ppm), suggesting that they all arise from residues in poorly structured, highly dynamic regions of the protein.

Previously, we developed an efficient protocol for the identification and optimization of peptide ligands from phage display libraries, giving rise to the peptide IGF-F1-1 that binds to IGF-I with an IC_{50} of ∼1 μM (*25*). Addition of IGF-F1-1 leads to a large and dramatic improvement in the quality of the NMR ^{1}H−^{15}N HSQC spectrum of IGF-I. Once saturating levels of peptide are added, over 60 backbone resonances are readily observed, and all peaks are of a much more uniform line width (supplementary Figure 1B). Addition of the peptide also leads to a vast improvement in the quality of other NMR spectra; representative regions of 2D ^{1}H NOESY spectra are shown in Figure 2. In the absence of peptide, the broad resonances and aggregated nature of IGF-I give rise to streaks rather than discrete peaks in the NOESY spectrum, particularly in the aromatic to methyl region.

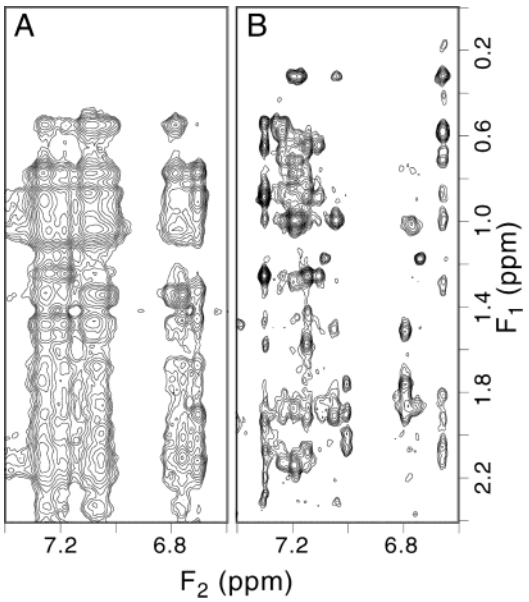

FIGURE 2: Representative sections of the ^{1}H NOESY spectrum of IGF-I in the absence (A) and presence (B) of peptide IGF-F1-1. Both samples contained 1 mM IGF-I in H_2O solution at pH 5.0 and were acquired under identical conditions (40 °C, 600 MHz spectrometer, τ_m = 100 ms, total acquisition time ∼18 h) and processed with the same apodization in F_1 and F_2.

However, the improved line widths in the complex allow many discrete and resolved peaks to be observed, suggesting that useful, high-resolution structural information can be obtained for the IGF-I peptide complex. Although the majority of the NMR spectra of IGF-I in complex with IGF-F1-1 were collected at pH 5.1 (see below), ^{1}H and ^{1}H−^{15}N HSQC spectra were collected over the pH range 5.1−7.3. The dispersion of ^{1}H resonances within the methyl envelope remained identical, and resonance positions of backbone ^{1}H−^{15}N groups did not change significantly. Thus, the conformation determined for IGF-I in the presence of IGF-F1-1 at pH 5.1 is maintained at neutral pH.

A series of spectra were obtained on complexes containing either uniformly ^{15}N- or ^{15}N/^{15}C-labeled IGF-I to allow structural characterization of the complex. Virtually all of the IGF-I side chain and backbone resonances could be assigned from the triple resonance spectra (supplementary Table 1), and many intra-IGF-I NOEs were assigned and used to generate 771 distance restraints. The observed medium-range NOEs suggest that the one B-region and two A-region helices observed in previous low- and high-resolution IGF-I structures are still present in the IGF-F1-1 complex (*14*, *15*, *17*, *19*−*21*). Moreover, the lack of medium- and long-range restraints for residues in the C- and D-regions (residues 30−41 and 63−70, respectively) indicates that they are disordered and probably dynamic in the complex. Indeed, from the assignments obtained for the complex, it is clear that the C- and D-regions give rise to the intense resonances observed in the ^{1}H−^{15}N HSQC spectrum of free IGF-I, presumably because they also exhibit high mobility in the absence of peptide.

Only partial resonance assignments for IGF-F1-1 could be obtained in the complex using ^{14}N/^{12}C-filtered experiments. We attribute this to exchange broadening of peptide resonances in the complex. Despite this, 47 intrapeptide distance restraints were identified, all of which were fully

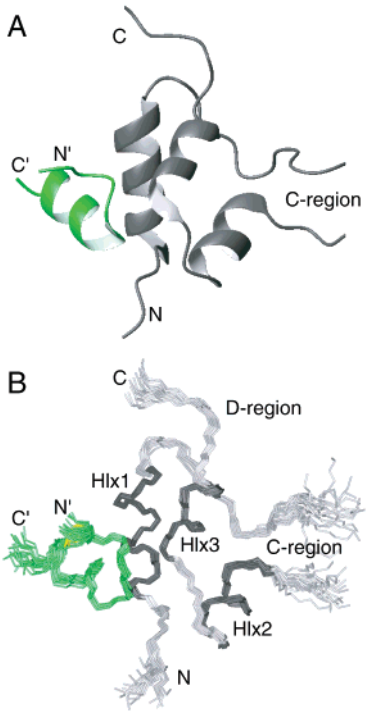

FIGURE 3: (A) Ribbon view of the complex formed between IGF-I (gray) and IGF-F1-1 (green). (B) Ensemble of 20 structures of IGF-1 (light gray loops and dark gray helices) bound to IGF-F1-1 (green). The structures were overlaid to the mean structure using N, Cα, and C atoms of residues 3−27, 41−61, and 2′−14′. For clarity, part of the disordered C-region (31−38) and the D-region (67−70) are omitted.

consistent with the structure determined for free IGF-F1-1 (*25*). Thus, during the structure calculation of the complex, the restraints based on the observable NOEs were supplemented by dihedral angle restraints within IGF-F1-1 to enforce a backbone conformation similar to that seen in the free state. Despite the broad nature of some peptide resonances, the 3D ω_1-^{12}C-filtered, ω_2-^{13}C-edited NOESY containing intermolecular contact information was of high quality (supplementary Figure 2). The majority of the intermolecular NOEs identified in this NOESY experiment could be assigned, giving rise to 87 intermolecular restraints to define the relative orientations of peptide and IGF-I in the complex.

Description of the Structure. The complex of IGF-I bound to IGF-F1-1 was determined on the basis of over 1000 experimentally determined restraints (Table 1). The final ensemble of structures agrees very well with the input restraints (no distance restraint violations >0.11 Å or dihedral angle restraint violations >1.5°) and is of good stereochemical quality (Table 1), as judged by the program PROCHECK (*37*). The secondary structure of IGF-I in complex with IGF-F1-1 is dominated by three helices (Figure 3A). The B-region α-helix (Gly 7−Cys18) is preceded by a short extended region (residues Glu3−Cys6) and followed by a type II′ reverse turn (Gly19−Asp20) that directs an extended region (Gly22−Tyr24) along the side of the B-region helix. A reverse-turn (Phe25−Asn26), type VIII in most members of the ensemble, orients the C-region Ω-loop (Gly30−Thr41) away from the helical core. The C-region merges into the A-region with an extended section (Thr41−Gly42) that is in proximity to the B-region type VIII turn. The two A-region helices (Ile43−Cys48, Leu54−Tyr60) are arranged approximately antiparallel to one another and are connected by another type VIII reverse turn (Phe49−Arg50) and an extended region (Ser51−Asp53). Although the first of these helices adopts a regular α conformation, the second is predominantly 3_{10} in character with *i*, *i* + 3 hydrogen bonds observed for most members of the ensemble in this region. The final residues of the A-region bend away from the B-region and lead into the disordered D-region C-terminus. The N-terminal two residues and the entire C- and D-regions are poorly defined by the NMR data, suggesting high mobility of these regions. However, the core B- and A-regions of IGF-I in the IGF-F1-1 complex are well-defined by the NMR data with a mean RMSD of backbone atoms to the mean structure of 0.27 ± 0.05 Å (Table 1, Figure 3B).

As in the case of the free peptide (*25*), IGF-F1-1 in the complex contains a reverse turn (Cys3′−Phe4′; note that throughout this work, the peptide residues are designated by the primed suffix) and a C-terminal α-helix (Ser6′−Tyr15′; Figure 3A). The cluster of hydrophobic residues on one face of IGF-F1-1 provides the surface for interaction with IGF-I: of the ∼500 Å^2 solvent-accessible surface area on the peptide that is buried on complex formation, over 75% is contributed by the hydrophobic residues Phe4′, Val7′, Leu10′, and Met14′ (Figure 4A). This surface on the peptide contacts a similarly sized hydrophobic patch on IGF-I (550 ± 26 Å^2 of buried surface area) formed by helix 1 and helix 3, with ∼55% of the buried surface contributed by five hydrophobic residues (Leu5, Ala13, Phe16, Val17, and Leu54; Figure 4A). The peptide helix extends from the N-terminus of helix 3 to the middle of helix 1 and has the extended N-terminal region of IGF-I running along one edge (Figure 3A). The side chain of Leu54 forms a striking protrusion on the IGF-I surface that is "pinched" by the peptide side chains Phe4′, Val7′, and Leu10′ (Figure 4A). These peptide side chains fit into a series of depressions on the IGF-I surface formed by an outer ring of IGF-I residues that are either hydrophobic (Leu5, Phe16, Val17) or charged (Glu3, Glu9, Asp12, Arg55, Glu58). The latter residues contact the hydrophobic peptide side chains via their methylene groups and may also have favorable polar interactions with the peptide. Due to the absence of protons at the termini of these side chains (i.e., no NOE restraints) and absence of Coulombic terms in the simulated annealing force field (i.e., no nonbonded attractive forces), the precision of the charged side chains in the final ensemble is not always sufficient to identify specific interactions involving these groups. However, the general proximity of side chains suggests potential intermolecular charge−charge or charge−dipole interactions from Glu3 to Ser6′ O$^\gamma$H or Val7′ H^N, from Glu9 to Arg11′, from Asp12 to Tyr15′ O$^\eta$H, and from Arg55 to Phe4′ CO or Glu5′ CO^{2-} (Figure 4A).

Functional Analysis of Peptide Residues. To assess the importance of individual peptide residues in recognition of IGF-I, we employed both individual residue and shotgun alanine scans in phage-based formats (*28*, *35*). In the first, IGF-F1-1 was displayed on phage and each non-cysteine residue was individually replaced with alanine (or glycine in the case of the two alanine residues in the peptide), and the effect on IGF-I binding was measured in a competition phage ELISA (Table 2). This procedure allows a rapid assay of the functional effect of mutation at many sites within IGF-

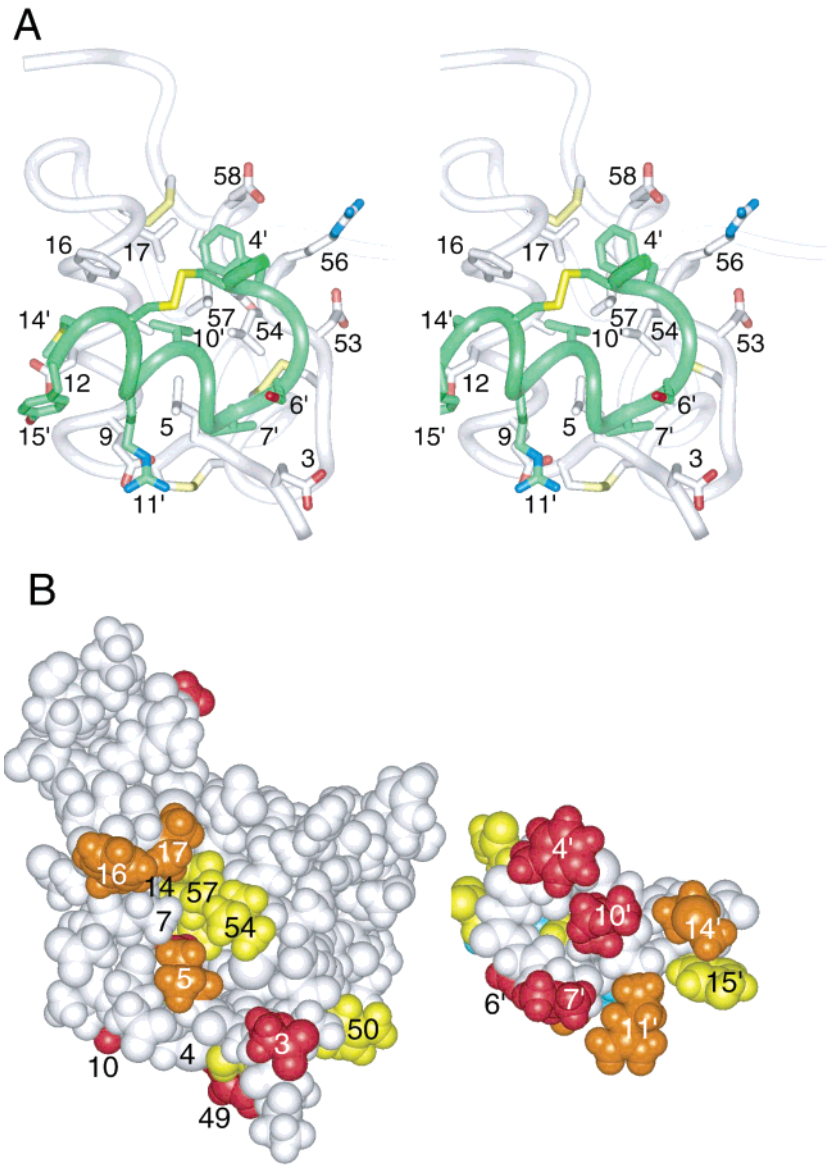

FIGURE 4: (A) Stereoview of interactions between IGF-F1-1 (green) and IGF-I (white). The backbone of each molecule is represented by a tube, while the interfacial side chains are depicted as sticks (nitrogen, oxygen, and sulfur atoms colored blue, red, and yellow, respectively). (B) Open book CPK representation of the interface between IGF-I and the IGF-F1-1. IGF-I is in the same orientation as in panel A whereas the peptide is rotated by 180° and translated in the plane of the page. Side chains on IGF-I are colored according to the effect that alanine substitution has on IGFBP-1 affinity: yellow, 2−5-fold effect; orange, 5−30-fold effect; red, >30-fold effect (*28*). The peptide side chains are colored according to the ratio of wild-type to alanine (glycine) identified in the combinatorial alanine (glycine) scan presented in Table 2: yellow, <10; orange, 10−100; red, >100.

F1-1 without the laborious synthesis, purification, and assay of individual peptide analogues. This analysis clearly demonstrates that the two N-terminal residues, the two alanine residues, and Arg12′ of IGF-F1-1 could be replaced by glycine or alanine without a significant decrease in binding to IGF-I. However, none of the other mutants gave rise to an IGF-I-dependent signal. The most likely explanation of this observation is that the majority of the mutations cause a reduction in the binding to IGF-I and that in these cases insufficient phage particles bind to the immobilized IGF-I to produce a signal. The importance of these residues in binding to IGF-I is not unexpected given their proximity in the solution structure of IGF-F1-1 (*25*). However, we cannot rule out the possibility that these mutations reduce the display of peptide on the surface of the phage particles and are therefore not directly reporting on a loss of function.

To circumvent the poor sensitivity in the alanine scan phage ELISA, a combinatorial or shotgun alanine scanning approach was adopted (*35*). In this approach, a phage library was constructed in which all 13 residues to be mutated were represented by degenerate codons that allowed either the wild-type amino acid or alanine (glycine in the case of wild-type alanine residues). Note that due to the limitations of the genetic code, some other mutations are also possible [see Table 2 (*35*)]. Those residues that are important for binding will be preferentially selected as the wild-type amino acid whereas residues that are not involved in binding will exhibit a more equal distribution of wild-type and alanine (or mutant amino acids). The diversity within such a 13-residue library is calculated to be 3.4×10^8, a value that is readily achieved experimentally (*34*). After several rounds of selection for binding to immobilized IGF-I, 192 positive clones were

Table 2: Alanine Scanning Mutagenesis Data for IGF-F1-1

residue	Ala scan[a] IC_{50}(mut/wt)	shotgun scan[c] (wt/mut) Ala	mut 1	mut 2
wt	1.0 (7.2 μM)			
1R	1.6	0.18 (8/44)	0.11 (8/71 G)	0.23 (8/35 P)
2N	1.3	0.4 (8/20)	0.07 (8/122D)	0.67 (8/12 T)
4F	nd	125 (125/1)	3.9 (125/32V)	>125 (125/0 S)
5E	nd	1.4 (95/65)		
6S	nd	>161 (161/0)		
7V	nd	165 (165/1)		
8A	2.8[b]		14.9 (149/10 G)	
9A	1.3[b]		7.4 (141/19 G)	
10L	nd	>159 (159/0)	>159 (159/0 P)	159 (159/1 V)
11R	nd	19 (153/16)	>153 (153/0 G)	>153 (153/0 P)
12R	1.5	0.08 (13/100)	0.28 (13/46 G)	>13 (13/0 P)
14M	nd	17.9 (143/8)	143 (143/0 T)	7.9 (143/18 V)
15Y	nd	2.3 (79/31)	1.5 (79/46 D)	5.5 (79/13 S)

[a] IC_{50} competition ELISA values are reported relative to that obtained for wild-type IGF-F1-1 displayed on the gene-3 coat protein (7.2 μM). nd indicates that the ELISA signal was not sufficiently strong to observe a decrease when soluble IGF-I was added. [b] Mutation to glycine. [c] Values indicate the ratio of clones containing the original IGF-F1-1 sequence divided by the number of clones containing the alanine mutant (or other mutant allowed by the particular nondegenerate codons used). In each case, 192 positive clones were analyzed, yielding 169 interpretable sequences (among the 169 sequences, as many as 11 were unreadable at certain positions).

sequenced, of which 169 gave rise to readable sequences for analysis (note that for some positions within the peptide, as few as 158 sequences could be read unambiguously). From the sequenced clones, residues Phe4′, Ser6′, Val7′, and Leu10′ are almost exclusively selected, and Arg11′ and Met14′ are highly represented, indicating that these residues are energetically important for binding to IGF-I (Table 2, Figure 4B). The very low ratio of wild-type to alanine observed for Arg1′ and Arg12′ may in part be due to a systematically low display of basic peptides, as noted previously (*38*), and suggests that Arg11′ may be even more important than the 19-fold ratio of arginine to alanine might otherwise indicate (Table 2). With the exception of Ser6′, all of the highly selected residues contribute to the hydrophobic and basic patch on the surface of IGF-F1-1 that contacts IGF-I, thereby explaining their energetic importance to binding (Figure 4). Since Ser6′ is not in direct contact with IGF-I, we rationalize the preference for serine by its ability to form an "N-cap" on the peptide helix that stabilizes the conformation necessary for binding.

The non-alanine mutations arising from the degenerate codon usage sheds further light on the requirements for tight binding. Thus, although substituting Phe4′ with alanine is not favorable, binding is observed with other hydrophobic residues at this site since ~20% of the sequences contained valine. Indeed, hydrophobic residues other than phenylalanine were observed at this position after the initial gene-8 phage selection and after the gene-3 optimization of the ligand (*25*). A similar result is seen in the data for Met14′ in that valine was observed during the original gene-8 selection and in ~11% of the sequences from the combinatorial alanine scan library. In contrast to this, Leu10′ is much less amenable to substitution, both in the combinatorial libraries described herein and the ligand optimization described previously (*25*). The single observation of a valine at this site (Table 2) suggests that binding to IGF-I is compromised either by less intimate contacts with the smaller side chain or through a destabilization of the peptide helix. No sequences were observed with proline or threonine substitutions for either Leu10′, Arg11′, Arg12′, or Met14′, a result consistent with the location of these residues within the helix and the low helical propensity of such substitutions. Interestingly, helix-breaking glycine substitutions are observed for Ala8′ (6%), Ala9′ (12%), and Arg12′ (29%), suggesting that the peptide can still bind to IGF-I in the presence of limited helix destabilization as long as direct contacts (hydrophobic or polar) with the protein are not removed.

DISCUSSION

The addition of the peptide IGF-F1-1 to IGF-I has the fortuitous effect of dramatically improving the quality of the IGF-I NMR resonances, giving rise to signals of a uniformly narrow nature (Figure 2). This has allowed for the first time the determination of a high-precision solution structure for much of IGF-I (Figure 3). The observed intermolecular NOEs are sufficient to identify with high certainty the interacting hydrophobic surfaces on IGF-I and the peptide. The resulting structures also suggest a number of specific polar interactions that may contribute to the binding energy of the peptide (Figure 4A). Two phage display-based selection methods have been used to demonstrate the importance of the hydrophobic and basic residues of IGF-F1-1 for tight association with IGF-I (Table 2, Figure 4B).

In our initial characterization of IGF-F1-1, the peptide was shown to inhibit association of IGF-I with both IGFBP-1 and IGFBP-3 with IC_{50} values of ~2 μM (*25*). Previously, several groups have used site-directed mutagenesis to identify residues within the B-region of IGF-I (helix 1 and the N-terminal strand) as being important for IGFBP association (Figures 1 and 4) (e.g., see refs *28*, *36*, and *39*−*41*). Such a contact region readily explains the observed inhibition by IGF-F1-1 since this is where the peptide binds (Figure 4B). Indeed, the N-terminal fragment of IGFBP-5 (mini-IGFBP-5) also binds to this region of IGF-I (see below), as observed in a recent cocrystal structure (*21*). Given the high sequence similarity among all six IGFBP family members in the N-terminal region (30−60%), it is likely that the N-terminal domains of all IGFBPs contact IGF-I in the same fashion, and thus none would be able to bind IGF-I in the presence of IGF-F1-1.

Although a structure has been published for the majority of the extracellular portion of the IGF-I receptor (*42*), the

complex that it forms with IGF-I has not thus far been amenable to crystallographic analysis. However, mutagenesis studies have shown that residues in the B-region (Tyr24), the C-region (Tyr31, Arg35, Arg36), and the A-region (Tyr60) of IGF-I are critical for tight receptor binding and signaling (summarized in Figure 1) (*43*, *44*). Even though the IGF-F1-1 binding site is not proximal to these regions of IGF-I, the peptide is nonetheless able to inhibit downstream phosphorylation events mediated by IGF-I binding to both I-R and IGF-R with an IC_{50} of $\sim$5 μM (*25*). Potentially, IGF-F1-1 might induce IGF-I to adopt a conformation that is unable to bind to receptor. However, we deem this possibility unlikely since the known receptor contact sites on IGF-I either are highly mobile in the presence of the peptide or else are in very similar conformations in all of the high-resolution IGF-I structures published to date (see below).

In regard to the receptor inhibition by IGF-F1-1, it is worth considering in more detail the proposed mechanism of signaling in the insulin system. Binding studies and kinetic measurements conducted on a large number of insulin analogues have led to a model where the initially formed complex is mediated by contacts from the "dimerization" face of insulin to only one of the α chains in the $\alpha_2\beta_2$ I-R heterotetramer. After this, a conformational change in the receptor is hypothesized that allows contacts to form between the second α chain and a different region on the same insulin molecule, leading to the high-affinity, signaling complex (*1*, *3*). Mutational evidence suggests that this second site on insulin is situated at the so-called "hexamer" interface (see refs 1−3) and comprises residues in the middle of helix 1 and the N-terminus of helix 3. Within IGF-I, at least some of the residues demonstrated to be important for IGF-R binding are in locations equivalent to the dimerization surface of insulin (Figure 1), suggesting that IGF-I and insulin share some similarity in the receptor recognition process, a conclusion also drawn from insulin-IGF chimera studies (*1*, *45*). More interestingly, residues in IGF-I that are equivalent to the insulin hexamer interface include Leu54 and Phe16, residues that are part of the IGF-F1-1 binding site. The negative cooperativity observed in insulin binding to its receptor has been attributed to "second site" interactions through the hexamer interface, and such cooperativity is also observed for IGF-I (*1*, *3*, *4*). Thus, if the binding model proposed for insulin is relevant to IGF-I, the perturbation of the second site on IGF-I by IGF-F1-1 might explain the inhibition of signaling by the peptide. We note that there are conflicting reports on the role of Phe16 in IGF-R binding: substitution with alanine causes a 50-fold reduction in binding in one report (*40*) whereas a separate study indicated that mutation to serine had no effect on receptor binding (*39*).

The present solution structure for IGF-I contains the same elements of regular secondary structure observed in the poorly defined structures reported previously for IGF-I in solution (*14*−*17*). A detailed structural comparison to these earlier IGF-I structures is not informative due to their low coordinate precision. However, three crystallographic studies have been reported recently that have yielded high-resolution structures of IGF-I (*19*−*21*). A common theme in these crystallographic studies and the present solution study is that high quality structures are *only* obtained in the presence of an ancillary molecule bound to IGF-I (either detergent, IGFBP fragment, or peptide). Moreover, in all cases the ligands bind to the same hydrophobic patch between the N-terminus of the B-region and helix 3 in the A-region. This hydrophobic patch surrounded by polar and charged residues apparently makes it an attractive binding site for many ligands. Targeting of a physiological protein−protein interaction site by the phage display process is not uncommon (e.g., see refs *46* and *47*), suggesting that such sites have specific qualities, either physiochemical or geometric, that lend themself to binding of both cognate and phage display-derived ligands.

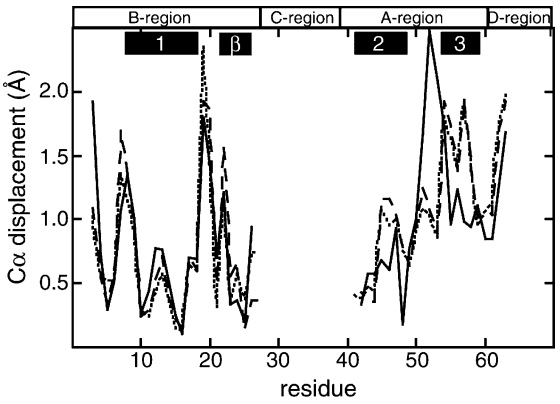

FIGURE 5: Displacement of Cα atoms of IGF-1 in the deoxy big CHAPS (dotted line), SB12 (dashed line), and mini-IGFBP-5 (solid line) complexes from the mean structure of the IGF-F1-1 complex. All structures were initially aligned using the backbone heavy atoms of the well-defined regions of IGF in the IGF-F1-1 complex (residues 3−27 and 42−63). The locations of the three helices, the short B-region β-strand, and the four regions of IGF-I are indicated by the boxes above the graph. Note that data are omitted for the ill-defined regions of the NMR ensemble (the Cα displacements for these residues are in the range 3−10 Å).

Structural comparison of IGF-I in the different complexes was initiated by analysis of Cα displacement plots following superposition of the structures using those residues that were well-defined in the NMR ensemble (3−26 and 42−63). Such a plot indicates that most of helix 1, the B-region β-strand, and helix 2 are quite similar in all structures (Cα displacement $\leq$1 Å; Figure 5). In particular, there is a very close correspondence in the side chain packing between the B-region helix and the B-region C-terminal strand (Figure 6A). However, there are significant differences in the loop immediately following helix 1 and the C-terminal half of the A-region (Figure 5). The Cα displacement analysis also reveals that the conformations of detergent-bound IGF-I all deviate from the peptide-bound form in the same manner regardless of detergent type and that a different structural perturbation is evident from the Cα displacements of IGF-I in the presence of mini-IGFBP-5. We hypothesize that structural perturbation within helix 3 and changes in its location relative to the B-region are necessary to accommodate the different ligands. Thus, helix 3 contains an N-terminal turn of 3_{10} helix and a C-terminal turn of regular α-helix in the detergent complexes but is entirely 3_{10} helix in the peptide complex. This has the effect of markedly changing the location of Leu54, a change that is necessary for the interactions with each ligand (Figure 6B). The conformation of helix 3 in the mini-IGFBP-5 complex is somewhat intermediate between the detergent and peptide-bound conformations (Figure 6C; Figure 7). Interestingly,

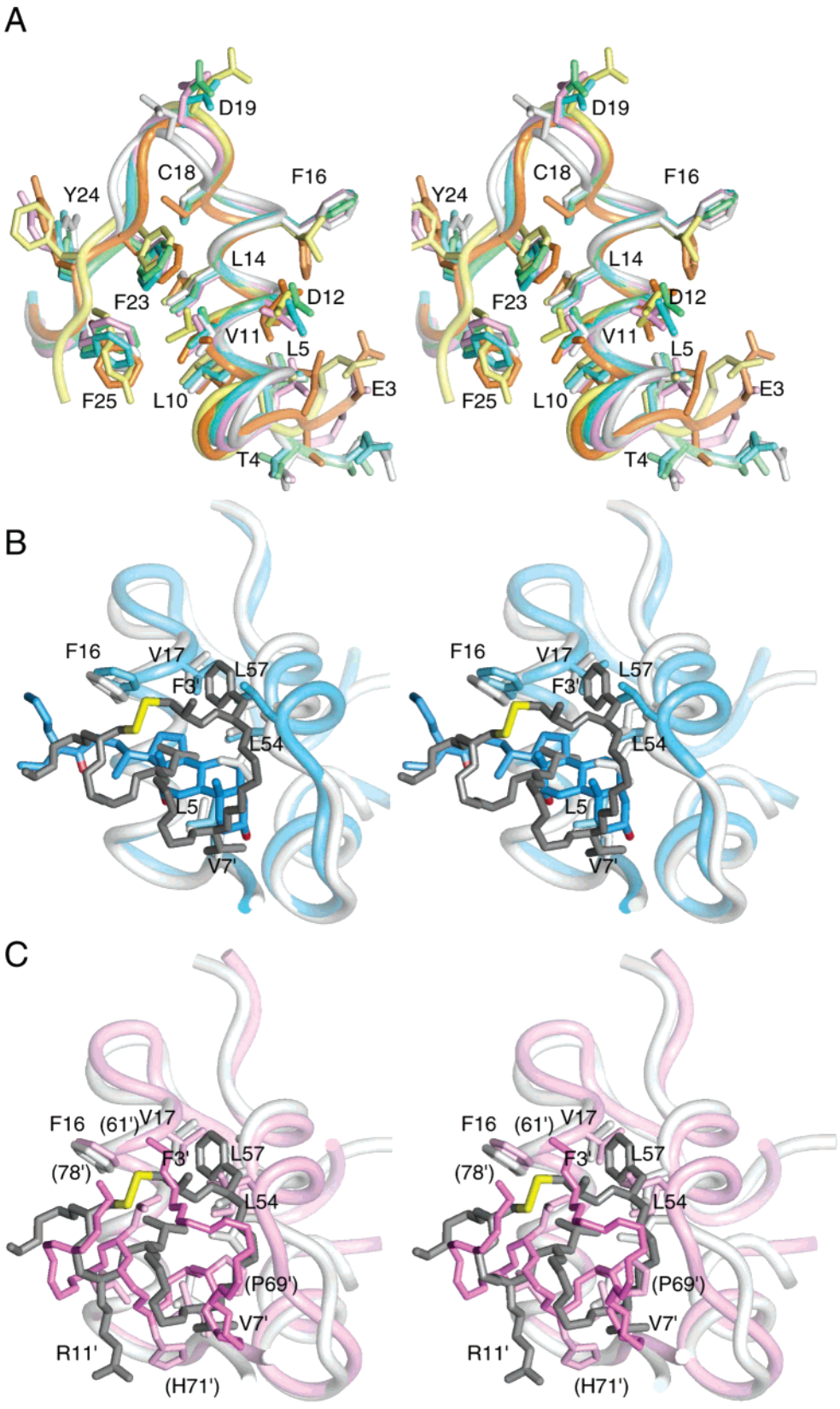

FIGURE 6: (A) Similarity in the B-region of IGF-I, IGF-II, and insulin. Backbone ribbons are depicted for residues 3−27 of IGF-I (B4−B28 of insulin) in the following complexes: IGF-F1-1 (white); deoxy big CHAPS (blue; PDB Accession Code 1imx) and SB12 (green; 1gzr); mini-IGFBP-5 (pink; 1h59). IGF-II (orange; 1igl model 7) and the 4-Zn^{2+} form of insulin (yellow; 4ins) are also shown. Some of the side chains are omitted for clarity. This orientation is rotated by ~90° about a vertical axis compared to that of Figure 4 (the peptide binding site is on the right in this view). (B) Differences in the helix 3 conformation for the peptide complex (IGF-I, white; IGF-F1-1, gray) and the deoxy big CHAPS complex (blue). IGF-I is in approximately the same orientation as in Figure 4. (C) Differences in helix 3 orientation for the peptide complex (IGF-I, white; IGF-F1-1, gray) and the mini-IGFBP-5 complex (IGF-I, light pink; mini-IGFBP-5, dark pink). IGF-I is in approximately the same orientation as in Figure 4.

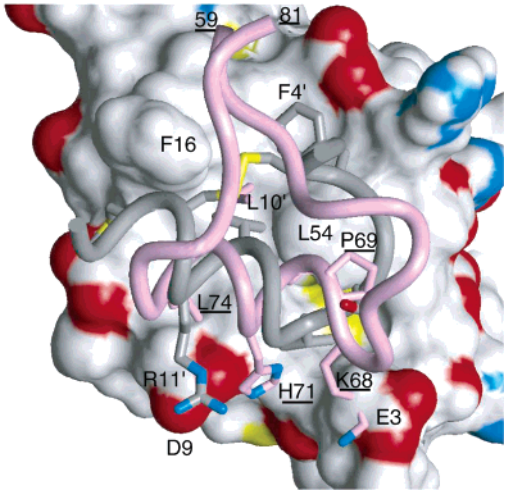

FIGURE 7: Comparisons of helix location of mini-IGFBP-5 (pink) and IGF-F1-1 (gray) when bound to IGF-I. IGF-I in the IGF-F1-1 complex is depicted as a white Connolly surface, with nitrogen, oxygen, and sulfur atoms colored blue, red, and yellow, respectively. Several important side chains at the interfaces are depicted (underlined labels refer to mini-IGFBP-5; primed numbers indicate residues in IGF-F1-1).

although both IGF-F1-1 and mini-IGFBP-5 use the hydrophobic face of a helix to bind to IGF-I, the helical backbones are offset by ~2 Å due to the capping mode and loop conformation prior to each helix (Figures 6C and 7). Thus, although the same patch of hydrophobic residues (Leu5, Phe16, Leu54, Leu57) is central to all of the complexes, there is a change in how these residues are used by IGF-I to recognize peptide, detergent, or binding protein.

Outside of the well-ordered B- and A-regions, there is less structural conservation between the different IGF-I complexes. Notably, most of the C-region is not well ordered in the IGF-F1-1 complex, and previous solution studies have shown that fast internal motions are responsible for the low restraint density of these residues (*17*). This, combined with the wide variation in loop location among the various crystal structures, suggests that although electron density is observed in this region, the exact location of the loop relative to the rest of IGF-I is probably influenced by the crystalline environment. A common feature in solution (*14*) and the crystalline state (*19*−*21*) is the presence of a reverse turn proximal to Tyr31. Although this conformation may indeed be necessary for recognition of Tyr31 by IGF-R (*43*), the variability elsewhere in the C-region limits the usefulness of this information since the precise location of Tyr31 with respect to other IGF-R contact residues is not well-defined. Furthermore, proteolytic clipping within the C-region was observed in one of the detergent complex studies, and this modification may have aided crystallization and induced subtle changes in the C-loop conformation (*20*). A detailed analysis of the through-bond scalar couplings and through-space sequential NOE patterns unequivocally demonstrates that the C-region in the present study is intact.

Inspection of structures determined for IGF-II (*18*) and insulin (*48*) indicates that the interactions between the B-region helix and B-region C-terminal strand are very similar to that seen in the IGF-I complexes described above (Figure 6A). Differences in the A-region conformation between IGF-I and IGF-II have been noted previously (*18*), and comparison with the structure determined for IGF-I herein also indicates a shift in the relative orientation of the B- and A-regions (the A-region of IGF-II is packed closer to the B-region, with a slight shift of helix 2 toward the N-terminus of the B-region). As with IGF-I, the poor quality NMR spectra of insulin have been attributed to a combination of self-association and internal dynamics. Although the exact origin of the internal dynamics in solution has been a matter of discussion (*49*−*52*), comparison of the structure of insulin in various crystal forms suggests that there exist a number of low-energy conformations of the A-chain that are able to pack against a conserved B-chain conformation (*53*). Thus, the presence of a structurally conserved B-region with conformational variability of the A-region residues appears to be a common theme for members of the insulin superfamily.

The current structural analysis of the IGF-F1-1 complex has identified a steric mechanism for the inhibition of association between IGF-I and IGFBP. Moreover, the structure also suggests a mechanism for the inhibition of receptor activation by the peptide that is in agreement with a two-site model of IGF-I binding to its receptor. Comparison of the conformations observed in the present study with those in the three reported crystal structures reveals a range of conformational variability within IGF-I. Thus, the majority of the B-region is relatively fixed in all four ligand-bound states, suggesting that the helix-strand packing forms a stable folding motif, a motif that is also utilized in IGF-II and insulin. In contrast to this, parts of the A-region appear to be quite malleable and adopt a range of conformations that are dependent on the particular ligand present in the complex. This observation may have some relevance to the biophysical properties of IGF-I. We propose that, in the absence of ligand, the A-region of IGF-I undergoes facile exchange between conformations on a time scale sufficiently slow to cause broadening of NMR resonances ("intermediate" chemical exchange). When combined with the fast internal dynamics of the C-loop, these motions also preclude the crystallization of IGF-I in a rigid lattice suitable for X-ray diffraction studies. Indeed, mobility within helix 2 and helix 3 has already been postulated on the basis of NMR relaxation measurements (*17*). Perhaps more importantly, the conformational variability of the A-region has functional significance. By containing regions of structural heterogeneity, this relatively small protein is not only able to bind tightly to nine physiological ligands (six different binding proteins and several signaling receptors) but can also associate with two classes of detergent and the phage-derived peptides described herein.

ACKNOWLEDGMENT

We thank Drs. Henry Lowman, Samantha Lien, Tim Stewart, Felix Vajdos, Mark Ultsch, and Ron Siddell for many fruitful discussions during the course of this work. We are also grateful to Drs. Wayne Fairbrother and Borlan Pan for assistance with the acquisition of the triple resonance spectra.

SUPPORTING INFORMATION AVAILABLE

Two figures showing regions of $^{1}H-^{15}N$ HSQC spectra of IGF-I with and without peptide and sections of the ω_1-^{12}C-filtered, ω_2-^{13}C-edited NOESY spectrum showing inter-

molecular NOEs to Leu54 and one table summarizing the NMR spectra acquired. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. De Meyts, P. (1994) *Diabetologia 37* (Suppl. 2), S135−S148.
2. De Meyts, P., Wallach, B., Christoffersen, C. T., Ursø, B., Grønskov, K., Latus, L.-J., Yakushiji, F., Ilondo, M. M., and Shymko, R. M. (1994) *Horm. Res. 42*, 152−169.
3. Schäffer, L. (1994) *Eur. J. Biochem. 227*, 1127−1132.
4. Christoffersen, C. T., Bornfeldt, K. E., Rotella, C. M., Gonzales, N., Vissing, H., Shymko, R. M., ten Noeve, J., Groffen, J., Heisterkamp, N., and de Meyts, P. (1994) *Endocrinology 135*, 472−475.
5. Ballard, F. J., Baxter, R. C., Binoux, M., Clemmons, D. R., Drop, S., Hall, K., Hintz, R., Rechler, M., Rutanen, E., and Schwander, J. (1989) *Acta Endocrinol. 121*, 751−752.
6. Baxter, R. C., and Martin, J. (1989) *Prog. Growth Factor Res. 1*, 49−68.
7. Lewitt, M., Saunders, H., and Baxter, R. C. (1993) *Endocrinology 133*, 1797−1802.
8. Baxter, R. C. (2000) *Am. J. Physiol. Endocrinol. Metab. 278*, E967−E976.
9. Blundell, T. L., Dodson, G. G., Hodgkin, D., and Mercola, D. (1972) *Adv. Protein Chem. 26*, 279−402.
10. Dodson, E. J., Dodson, G. G., Hubbard, R. E., and Reynolds, C. D. (1983) *Biopolymers 22*, 281−291.
11. Brange, J., Ribel, U., Hansen, J. F., Dodson, G. G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., and Voigt, H. O. (1988) *Nature 333*, 679−682.
12. Weiss, M. A., Hua, Q.-X., Lynch, C. S., Frank, B. H., and Shoelson, S. E. (1991) *Biochemistry 30*, 7373−7389.
13. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.
14. Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991) *Biochemistry 30*, 5484−5491.
15. Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T., and Kobayashi, Y. (1993) *Int. J. Pept. Protein Res. 41*, 433−440.
16. De Wolf, E., Gill, R., Geddes, S., Pitts, J., Wollmer, A., and Grötzinger, J. (1996) *Protein Sci. 5*, 2193−2202.
17. Laajoki, L. G., Francis, G. L., Wallace, J. C., Carver, J. A., and Keniry, M. A. (2000) *J. Biol. Chem. 275*, 10009−10015.
18. Torres, A. M., Forbes, B. E., Aplin, S. E., Wallace, J. C., Francis, G. L., and Norton, R. S. (1995) *J. Mol. Biol. 248*, 385−401.
19. Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001) *Biochemistry 40*, 11022−11029.
20. Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenberg, J. P., de Bree, F. M., and Dauter, Z. (2002) *Biochemistry 41*, 9389−9397.
21. Zeslawski, W., Beisel, H. G., Kamionka, M., Kalus, W., Engh, R. A., Huber, R., and Lang, K. H., T A. (2001) *EMBO J. 20*, 3638−3644.
22. Lowman, H. B., Chen, Y. M., Skelton, N. J., Mortensen, D. L., Tomlinson, E. E., Sadick, M. D., Robinson, I. C. A., and Clark, R. G. (1998) *Biochemistry 37*, 8870−8878.
23. Skelton, N. J., Chen, Y. M., Dubree, N., Quan, C., Jackson, D. Y., Cochran, A. G., Zobel, K., Deshayes, K., Baca, M., Pisabarro, M. T., and Lowman, H. B. (2001) *Biochemistry 40*, 8487−8498.
24. Loddick, S. A., Liu, X.-J., Liu, Z.-X., Liu, C., Behan, D. P., Chalmers, D. C., Foster, A. C., Vale, W. W., Ling, N., and De Souze, E. D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 1894−1898.
25. Deshayes, K., Schaffer, M. L., Skelton, N. J., Nakamura, G. R., Kadkhodayan, S., and Sidhu, S. S. (2002) *Chem. Biol. 9*, 495−506.
26. Swartz, J. (1994) Method for producing polypeptides via bacterial fermentation, U.S. Patent No. 5342763.
27. Reilly, D., and Fairbrother, W. J. (1994) *J. Biomol. NMR 4*, 459−462.
28. Dubaquié, Y., and Lowman, H. B. (1999) *Biochemistry 38*, 6386−6396.
29. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1995) *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, New York.
30. Pan, B., Li, B., Russell, S. J., Tom, J. K., Cochran, A. G., and Fairbrother, W. J. (2002) *J. Mol. Biol. 316*, 769−787.
31. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
32. Lowman, H., Bass, S., Simpson, N., and Wells, J. (1991) *Biochemistry 30*, 10832−10838.
33. Kunkel, K. A., Bebenek, K., and McClary, J. (1991) *Methods Enzymol. 204*, 125−139.
34. Sidhu, S. S., Weiss, G. A., and Wells, J. A. (2000) *J. Mol. Biol. 296*, 487−495.
35. Weiss, G. A., Watanabe, C. K., Zhong, A., Goddard, A., and Sidhu, S. S. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 8950−8954.
36. Jansson, M., Andersson, G., Uhlén, M., Nilsson, B., and Kördel, J. (1998) *J. Biol. Chem. 273*, 24701−24707.
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
38. Peters, F., Schatz, P., Johnson, S., and Dower, W. (1994) *J. Bacteriol. 176*, 296−305.
39. Magee, B., Shooter, G. K., Wallace, J. C., and Francis, G. L. (1999) *Biochemistry 38*, 15863−15870.
40. Jansson, M., Uhlen, M., and Nilsson, B. (1997) *Biochemistry 36*, 4108−4117.
41. Clemmons, D. R., Dehoff, M. L., Busby, W. H., Bayne, M. L., and Cascieri, M. A. (1992) *Endocrinology 131*, 890−895.
42. Garrett, T., McKern, N., Lou, M., Frenkel, M., Bentley, J., Lovrecz, G., TC, E., Cosgrove, L., and Ward, C. (1998) *Nature 394*, 395−399.
43. Bayne, M. L., Applebaum, J., Chicci, G. G., Miller, R. E., and Cascieri, M. A. (1990) *J. Biol. Chem. 265*, 15648−15652.
44. Zhang, W., Gustafson, T. A., Rutter, W. J., and Johnson, J. D. (1994) *J. Biol. Chem. 269*, 10609−10613.
45. Cara, J. F., Mirmira, R. G., Nakagawa, S. H., and Tager, H. S. (1990) *J. Biol. Chem. 265*, 17820−17825.
46. Sidhu, S. S. (2000) *Curr. Opin. Biotechnol. 11*, 610−616.
47. Sidhu, S. S., Fairbrother, W. J., and Deshayes, K. (2003) *Chem. Biochem. 4*, 14−25.
48. Baker, E., Blundell, T., Cutfield, J., Cutfield, S., Dodson, E. J., Dodson, G., Hodgkin, D., Hubbard, R. E., Isaacs, N., and Reynolds, C. (1988) *Philos. Trans. R. Soc. London, Ser. B 319*, 369−456.
49. Hua, Q.-X., Kochoyan, M., and Weiss, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2379−2383.
50. Hua, Q.-X., Ladbury, J., and Weiss, M. (1993) *Biochemistry 32*, 1433−1442.
51. Ludvigsen, S., Roy, M., Thøgersen, H., and Kaarsholm, N. (1994) *Biochemistry 33*, 7998.
52. Olsen, H., Ludvigsen, S., and Kaarsholm, N. (1996) *Biochemistry 35*, 8836−8845.
53. Badger, J. (1992) *Biophys. J. 61*, 816−819.

BI034386C