Chemistry & Biology, Vol. 9, 495–505, April, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00129-1

Rapid Identification of Small Binding Motifs with High-Throughput Phage Display: Discovery of Peptidic Antagonists of IGF-1 Function

Kurt Deshayes,[1,3] Michelle L. Schaffer,[1] Nicholas J. Skelton,[1] Gerald R. Nakamura,[1] Saloumeh Kadkhodayan,[2] and Sachdev S. Sidhu[1,3]
[1]Department of Protein Engineering
[2]Department of Bioanalytical Research and Development
Genentech
1 DNA Way
South San Francisco, California 94080

Summary

A panel of 22 naïve peptide libraries was constructed in a polyvalent phage display format and sorted against insulin-like growth factor-1 (IGF-1). The libraries were pooled to achieve a total diversity of 4.4×10^{11}. After three rounds of selection, the majority of the phage clones bound specifically to IGF-1, with a disulfide-constrained CX_9C scaffold dominating the selection. Four monovalently displayed sub-libraries were designed on the basis of these conserved motifs. Sub-library maturation in a monovalent format yielded an antagonistic peptide that inhibited the interactions between IGF-1 and two cell-surface receptors and those between IGF-1 and two soluble IGF binding proteins with micromolar potency. NMR analysis revealed that the peptide is highly structured in the absence of IGF-1, and peptides that preorganize the binding elements were selected during the sorting.

Introduction

Exploration of the cellular machinery is leading to the discovery of complex signaling pathways that are often initiated by an extracellular ligand binding to a cell surface receptor [1]. By specifically targeting these interactions, researchers have been able to manipulate cellular events in an attempt to one day alleviate disease. Inevitably, controlling signal transduction with the specificity necessary for therapeutic utility requires understanding manifold protein-protein interactions at the molecular level. Antibodies, peptides, and small molecules are powerful tools for investigating protein-protein interactions, but their discovery and development can be labor intensive and time consuming. Over the past decade, combinatorial methods such as phage display have assumed a critical role in the discovery of molecules used for elucidating key ligand-receptor relationships (reviewed in [2]). Even with the large libraries accessible via phage display, a recurrent problem is that, as theoretical library size grows (i.e., as the size of the displayed peptides increases and more positions have to be randomized), representation of a significant proportion of its constituents in a real, physical library becomes essentially impossible. Perturbation of protein-protein interactions is particularly demanding because their binding surfaces often contain few distinguishing features [3], and therefore large, structurally diverse libraries are required for identifying potent ligands.

M13 phage have been widely used as a display tool for the optimization of loops within proteins [4] and for the identification of novel ligands [2]. M13 lends itself to this process because engineering the viral DNA can lead to the fusion of peptides or proteins of interest at the surface-exposed termini of several of its coat proteins. One can create naïve libraries of phage having many different molecules fused to the coat by randomizing certain codons within the synthetic oligonucleotides used to engineer the viral DNA. Once the phage library is enriched with active peptides by binding phage particles to immobilized target, DNA sequencing may be used to identify those coat-protein-fused molecules that are ligands to the target because of the link between phenotype (displayed molecule) and genotype (viral DNA).

An important factor in the selection process is the choice of coat protein for library display. Display on the major coat protein (protein-8, P8) gives polyvalent display, with as many as several hundred copies of the coat fusion protein present on each phage particle. This has the advantage of providing an avidity component to binding to the immobilized target and often facilitates the selection of ligands that would bind weakly as monomers. As a caveat, the avidity may confound measurement of the tightness of binding (for example, in a phage enzyme-linked immunosorbant assay [ELISA], see below), and enrichment of certain library members may not correlate directly with target affinity of the library monomer. Alternatively, the library members may be displayed on a minor coat protein (protein-3, P3). Only a few copies of this protein are present in each phage particle, and by suitable engineering, the display of library members can be limited to one per phage particle [5]. The monomeric nature of this display lends itself to selection for tight binding (for example, if one uses more stringent selection conditions) and for the quantitation of library members' affinity for the target in phage ELISA assays.

Various scaffolds have been used in previous discoveries of novel ligands from naïve phage-peptide libraries. Such scaffolds have included unconstrained sequences, conotoxin derivatives [6], and peptides structurally constrained through a combination of pairs of cysteine residues (potential disulfide bonds) or Gly-Pro motifs (potential reverse turns) [7–10]. Commonly, peptide-phage libraries have been generated by combining pools of phage displaying similar scaffolds (e.g., structurally constrained peptides kept separate from linear motifs), with each library being pursued independently (e.g., see [8]). Parallel manipulation of several libraries increases the labor involved but has the advantage that several scaffolds capable of binding may be obtained. The identification of multiple peptide classes able to bind to a target is advantageous in that it allows subsequent categorization on the basis of target affinity,

[3]Correspondence: deshayes@gene.com (K.D.) sidhu@gene.com (S.S.S.)

target specificity, or ligand characteristics (size, structure, solubility). This may not be possible in an "all-in-one" library because avidity effects or expression-biased phage growth during amplification in *E. coli* may lead to the domination of the selection process by one or a limited number of clones (e.g., see [9]). Because of the benefits of avidity, initial screening has routinely been done on P8 with many rounds of selection (3–6) to yield a small number of independent clones. At this point, peptide synthesis is used to make specific peptides and test for activity. This can be an expensive and time-consuming process, particularly if the initial ligands are of low affinity. Only once this has been accomplished are secondary phage libraries designed as P3 constructs to optimize the initial ligands.

Recent modifications to the phage display technique have led to significant gains in practically accessible sequence diversity [5]. These modifications include (1) increasing the size of peptide libraries to more than 10^{10} members; (2) combining several libraries so that several scaffolds can be investigated simultaneously; (3) mining sequence data from early rounds of sorting to obtain information with which to design a directed search of sequences space; and (4) developing a protocol for the accurate and rapid analysis of hundreds, or even thousands, of clones. Herein, we describe how we have incorporated these improvements into our ligand discovery process with the aim of decreasing the time, effort, and cost of obtaining peptides that bind to a particular target protein (Figure 1).

We originally cast a broad net in the naïve polyvalent sort by sampling an extremely diverse collection of sequences and scaffolds for potential binding motifs. The phage-displayed libraries were pooled and cycled through rounds of binding selection with immobilized target protein to select for members that bound specifically to the target. After a minimal number of selection cycles, phage binding ELISAs were used to identify individual clones that bound specifically to the target. The sequences of binding peptides were then deduced from the sequences of the cognate DNA encapsulated within the phage particles. Because an extremely diverse library pool had been subjected to a low-stringency selection process, we sought to obtain many positive clones that could be grouped into families on the basis of amino acid sequence similarity.

Second-generation libraries were displayed in a monovalent format on P3 to enable stringent selection and evolution of high-affinity binders through a combination of changes in backbone structure and/or interacting side chains. Furthermore, while peptides were displayed monovalently on phage, their affinities could be assayed directly [5], allowing us to continuously monitor the selection process to identify high-affinity binders. Once affinity had been optimized to a satisfactory level, peptide sequences were synthesized and subjected to detailed analysis of structure and function. This approach has several advantages, including (1) time saved because of sorting a single combined library; (2) reduced expression bias among the selectants because of limited rounds of sorting; (3) utilization of fast DNA sequencing of many positive clones; (4) identification and maturation of multiple scaffolds at one time; (5) rapid evaluation of selection or optimization by phage-ELISA methods; and (6) limited peptide synthesis only after high-affinity ligands have been identified.

The target investigated in this report is insulin-like growth factor (IGF-1), a 70 residue peptide hormone that regulates both mitogenic and metabolic functions, primarily via binding to the cell surface IGF-1 receptor (IGFR), an $\alpha_2\beta_2$ heterotetramer related to the insulin receptor (IR) (reviewed in [11]). IGF-1 activity is tightly regulated, with IGF-1 existing in vivo primarily bound to one of six soluble IGF binding proteins (IGFBPs) that regulate IGF-1 biological availability and plasma half-life by sequestering the hormone within high-affinity complexes [12–16].

IGF-1 has a broad range of biological activities, and maintaining proper levels of IGF-1 signaling is important for normal cellular function [17–19]. A significant challenge in developing therapeutics based upon IGF-1 is understanding the interplay of the interactions that control IGF-1 signaling. Evidence for specific epitopes on IGF-1 for the different receptors has been obtained through alanine scanning mutagenesis [20–22]. Additional information about ligand binding to IGF-1 is found in crystal structures of IGF-1 bound either to a fragment of IGFBP5 [23] or a detergent (deoxybigchaps) [24].

In this paper, we discuss the application of our refined phage display technology to the discovery of peptide antagonists of IGF-1. Such molecules would have considerable utility in unraveling IGF-1 function, e.g., by antagonizing its effects in cell-based assays. We describe the rapid progression from large naïve libraries to more focused libraries to active synthetic peptides. Library design and information gathering and sorting procedures are presented, along with in vivo potency and NMR structural data. Finally, in order to elucidate the mechanism by which the affinity maturation occurred, we compare the structural features of a potent peptide obtained from monovalent maturation to that of a low-affinity parent from the original naïve sort.

Results

Naïve Polyvalent Libraries

As part of the streamlined protocol used in the present study, we chose to minimize the amount of manual intervention by using a single, combined library containing 22 P8 polyvalently displayed peptide-phage libraries (Table 1) with a combined diversity of 4.4×10^{11}. These scaffolds included linear sequences, conotoxin derivatives [6], and peptides structurally constrained through the inclusion of pairs of cysteine residues and Gly-Pro motifs. Multiple rounds of sorting are commonly employed to enrich the fraction of binding to nonbinding clones. However, "over sorting" may lead to domination by preferentially expressed clones and a loss of sequence diversity. Thus, we employed an ELISA-based assay after every round of sorting to measure the proportion of clones capable of binding IGF-1. In this way, we maintained diversity among the selected sequences by halting the selection process as soon as the fraction of binding to nonbinding clones reached a predetermined level. The single-point ELISA for measuring IGF-1

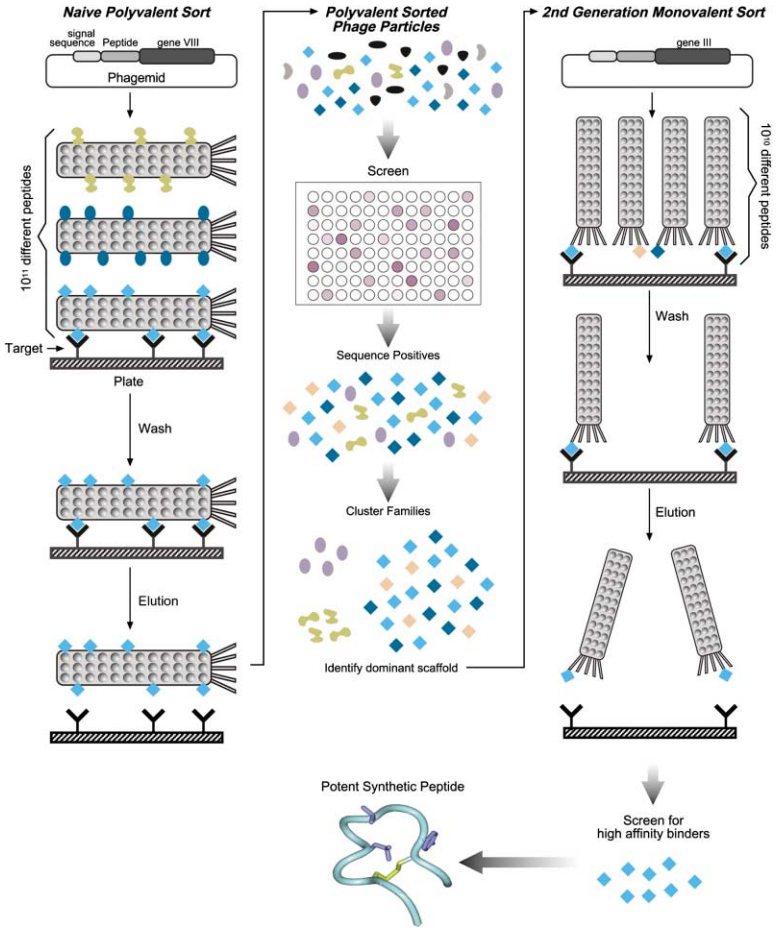

Figure 1. Schematic Diagram for the Polyvalent Naïve Selection and Monovalent Affinity Maturation of Binding Motifs

binding uses anti-phage antibodies to detect phage bound to immobilized IGF-1 (see Experimental Procedures). This procedure is readily adapted to a high-throughput format; hence, this additional step in the sorting protocol does not significantly increase the overall time of the sorting process.

After three rounds of sorting, more than 90% of the picked clones were positive for IGF-1 binding. Nonselective binding was ruled out because these clones did not bind to immobilized bovine serum albumin (BSA). Ninety-six colonies selected at this point were sequenced to identify the peptides responsible for IGF-1 binding. Gratifyingly, diversity was maintained at both the sequence and the scaffold levels. Representatives from six of the initial peptide sub-libraries were present in the 28 unique clones (Table 2). Within the most populous scaffold (CX_9C), there was considerable sequence diversity. These data suggest that sequence information from a single 96-well plate is capable of identifying preferred scaffolds and that, for at least some of these scaffolds, a variety of sequences capable of binding are observed (see Experimental Procedures). Given the efficiency of microtiter plate-based DNA sequencing, this amount of sequence data is readily obtained and thus does not increase the overall time required to obtain ligands to a given target. We note that sequencing of additional colonies may identify additional sequence information for other scaffolds and thereby permit a similar analysis to that described below for the highly populated CX_9C scaffold.

The CX_9C scaffold is most prevalent in the sequenced clones, and we focus on this family for the remainder of our analysis. Alignment of the 15 sequences on the basis of cysteine position reveals a number of conserved residues within the $X_2CX_9CX_2$ core. From this alignment, the sequences fall into two related families (Table 2). Family 1 has a consensus of serine at position 6, leucine at position 10, and methionine at position 14. Positions 7, 9, and 11 show a preponderance of two or three amino acid types, e.g., position 11 favors basic residues whereas position 7 shows a clustering of small, hydrophobic groups. Family 2 exhibits a consensus of aspartic acid at position 6, tyrosine at position 12, and tryptophan at position 14.

Monovalent Library Design and Selection

Usually, once a few unique clones have been identified from phage naïve peptide libraries, peptides would be synthesized, and further biochemical characterization would be performed. Information gleaned from these studies would form the basis of second-generation phage library design that would be used in affinity maturation of the initial ligand. In the present streamlined analysis of IGF-1 binding peptides, we chose not to perform costly and time-consuming peptide synthesis at this stage. Instead, we were able to use the sequence

Table 1. Naïve Polyvalent Peptide-Phage Libraries Sorted for Binding to IGF-1

Library[a]	Design[b]	Diversity[c] ($\times 10^{10}$)
1	X_8	2.6
2	X_{20}	1.2
3	$CX_6CX_6CCX_3CX_6C$[d]	1.6
4	CCX_3CX_6C[d]	1.7
5	$CCX_5CX_4CX_4CC$[d]	1.6
6	$CXCX_7CX_3CX_6$[d]	1.5
7	$X_4CX_2GPX_4CX_4$	2.0
8	CX_2GPX_4C	2.5
9	$X_7CX_4CX_7$	2.5
10	$X_7CX_5CX_6$	1.4
11	$X_6CX_6CX_6$	2.5
12	$X_6CX_7CX_5$	2.1
13	$X_5CX_8CX_5$	1.9
14	$X_5CX_9CX_4$	2.0
15	$X_4CX_{10}CX_4$	2.5
16	$X_3CX_4CX_2$	2.1
17	$X_2CX_5CX_2$	2.2
18	$X_2CX_6CX_2$	1.5
19	$X_2CX_7CX_2$	2.1
20	$X_2CX_8CX_2$	2.1
21	$X_2CX_9CX_2$	2.2
22	$X_2CX_{10}CX_2$	2.4
Total		44.2

[a] Libraries were displayed on gene-8, under the control of the P_{tac} promoter.
[b] Fixed residues are denoted by the single-letter amino acid code, and "X" denotes a randomized position where an NNK degenerate DNA codon encodes all 20 natural amino acids.
[c] For sorting, the libraries were pooled to achieve a total diversity of 4.4×10^{11}.
[d] Conotoxin derivatives.

Table 2. IGF-1 Binding Peptides Selected from Naïve Libraries

Scaffold[b]	Peptide Sequence[a]
X_{20}	ASQTPWPYSILFGEWWNAGF
	EAGAESRGWLQARCGELLGV
	HWDWTGGYWWIGREPWKEAA
	RLNAEGLRMGWGYMVWHWLS
$CXCX_7CX_3CX_6$	CECGKVSSRGCEKLCWLVSYM
CX_6C	GAWAWLCEQREEWCGWMLGT
	YDWVEACQKWPVLCMDSTMY
CX_7	GIREELCDKGLHKMCFREVR
CX_8C	DAMDCVVGPEWRKCFLEGG
	SGTACRWGPSWLLCSLASS
	GEGPECDLRQWGNLCSHWET
	LSSEECWEALKWQGCLMSER
	SFCEFNDWWPTCLV
CX_9C[c]	VKDECLMSVEALKNCMGLVS
	IDCLDSVEALKQCMY
	VESECLLSLPNLRRCMMDRL
Family 1	HPDKCFADVRALQECMESVR
	VEEKCYESITALRHCMQAMQ
	QIPAGCYESVQSLLECVQSA
	VMDQCFESYAEMRKCMLDGS
	EVARCVVDAGGTWYCWAEMA
	IECWQDLQGTRLCWE
	GESTCVTDLERVEYCWDEKS
Family 2	STYSCIRDMGWAVYCWETTL
	GVETCYSDAMNTQYCWTTEL
	REVKCMKDLSGHEYCWAEPR
	GEAAECAYDSLGMAYCYAKE
	TAGIECAYDKHLDQYCWWKE

The libraries shown in Table 1 were pooled and cycled through three rounds of selection for IGF-1 binding.
[a] Underlined residues were fixed as cysteines in the library constructions.
[b] Sequences were grouped into scaffold classes on the basis of cysteine spacing.
[c] Most sequences were from the CX_9C scaffold, and these could be divided into two families on the basis of sequence conservation (shown in bold text).

similarity information obtained from the initial polyvalent selection to guide the design of second-generation phage libraries. In order to keep the theoretical diversity within the limits of the library size and still fully explore the sequence space suggested by the similarity alignment, we constructed two libraries for each family with a combination of "hard" randomization (allowing for the expression of any amino acid), and "tailored" randomization (allowing for the expression of only a particular subset of amino acids) (Table 3).

Different tailored codons were used at the same positions within each sub-library, allowing sequence variability while limiting overall library size. For example, in position two of family 1, library F1-A uses the VVK codon, whereas F1-B uses the GAS codon (Table 3). Use of the VVK codon excludes uncharged hydrophobic residues in this position, and use of the GAS codon takes advantage of residue conservation by allowing incorporation of only the negatively charged aspartate and glutamate residues. For both F1 libraries, the first and last residues were hard randomized with the NNK codon, which encodes all twenty natural amino acids. In the design of both the F1 and F2 libraries, the disulfide constraint was left intact. Most of the residues within the loop between the cysteines showed some level of conservation, allowing us to design library F1-A with more unique codons and library F1-B with more degenerate codons. A smaller hydrophobic motif was recognized within the family 2 sequences (Table 2). The residues flanking the second cysteine displayed conservation as large hydrophobes (tyrosine and tryptophan), and positions 4 and 7 had a propensity toward small hydrophobic residues (Table 3). The tyrosine and tryptophan were held constant, and other positions took advantage of tailored codons encoding either charged residues, small residues, or nonaromatic hydrophobic residues. The NNK codon was used more frequently, giving larger theoretical diversities for the family 2 sub-libraries. However, the theoretical diversity was overrepresented in each case, with libraries containing from 2.0×10^{10}–2.6×10^{10} unique members (see Table 3). For binding selection, the A and B sub-libraries were combined for each family; families 1 and 2 were sorted separately.

Peptides were displayed at the N terminus of human growth hormone (hGH) fused to the C-terminal domain of the gene-3 minor coat protein. This ensures that the peptides are displayed in a stringent monovalent format because hGH is displayed at low levels on M13 phage [4]. An additional benefit of this display format is that it allows one to test for monovalent display. After a sort against IGF-1 for primary binding function, a second

Table 3. Second-Generation Library Designs for Affinity Maturation of IGF-1 Binding Peptides

Library[a]	DNA Codons[b] *Encoded Amino Acids*															Diversity[c]	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Theoretical	Actual
IGF-F1-A	NNK *All 20*	VVK *PHQ RTNK SADE G*	TGC *C*	YWC *FYLH*	GAS *DE*	AGC *S*	GTC *V*	VVK *PHQ RTNK SADE G*	GCT *A*	CTC *L*	ARG *KR*	VVK *PHQ RTNK SADE G*	TGT *C*	ATG *M*	NNK *All 20*	1.5×10^{6}	2.0×10^{10}
IGF-F1-B	NNK *All 20*	GAS *DE*	TGC *C*	YWC *FYLH*	RNK *IMTN KSRV ADE G*	AGC *S*	GTC *V*	GAG *E*	GCT *A*	CTC *L*	ARG *KR*	NNK *All 20*	TGT *C*	ATG *M*	NNK *All 20*	1.1×10^{7}	2.2×10^{10}
IGF-F2-A	NNK *All 20*	RVK *TNKS RAD EG*	TGC *C*	RYK *IMVT A*	NNK *All 20*	GAC *D*	KYG *LSVA*	NNK *All 20*	GGT *G*	NNK *All 20*	NNK *All 20*	TAC *Y*	TGT *C*	TGG *W*	GMK *ADE*	1.7×10^{9}	2.6×10^{10}
IGF-F2-B	NNK *All 20*	ARG *KR*	TGC *C*	RYK *IMVT A*	NNK *All 20*	GAC *D*	KYG *LSVA*	RGT *GS*	GGT *G*	NNK *All 20*	NNK *All 20*	TAC *Y*	TGT *C*	TGG *W*	GMK *ADE*	3.8×10^{7}	2.0×10^{10}

[a]To ensure monovalent display, we linked the libraries to hGH, which in turn was fused to the C-terminal domain of P3.
[b]Each library encoded peptides containing 15 residues. Tailored DNA codons were used to vary the level of amino acid diversity at each of the 15 positions within each library. DNA codons are shown in bold text, and the encoded amino acids are shown below in italics. Fixed, moderately randomized, and hard-randomized positions are shown in black, green and red, respectively.
[c]For each library design, the theoretical diversity (the number of different amino acid combinations encoded by the degenerate codons) was exceeded by the actual diversity of the constructed library.

sort against an anti-hGH antibody is conducted to ensure that only peptides displayed on the N terminus of hGH have been selected. After four rounds of panning, ELISA analysis of randomly selected clones showed the presence of selectants that bound to IGF-1 and anti-hGH antibody, with the majority of the clones showing a strong response. None of the clones bound to BSA, indicating we had generated a pool of monovalent peptides that bound selectively to IGF-1.

An additional selection round was performed under more stringent conditions to remove the weaker binding clones from the pool. Phage were bound to a low concentration (1 nM) of biotinylated IGF-1 in solution; rapid capture with streptavidin-coated magnetic followed. Using a low concentration of target protein increased the selection stringency to allow capture of only the highest affinity phage clones. After this final round of sorting, all the clones tested in the ELISA assay were positive for binding to both IGF-1 and anti-hGH antibody.

Assessment of Binding Affinity with Competition ELISA

The low level of peptide display with the monovalent hGH-fusion format allows the assessment of peptide affinity via a competition analysis while the peptide is still attached to the phage particle [5]. This procedure gives information about potency and allows us to rapidly identify potential leads for synthesis and further characterization. The clones were examined via an ELISA in which soluble IGF-1 inhibited phage-displayed peptides from binding to immobilized IGF-1 (Experimental Procedures); anti-M13 phage antibody HRP conjugate is used to detect the phage bound to immobilized IGF-1, and an IC_{50} value is determined. Importantly, this assay format allows the affinity of any hGH-peptide fusion to be determined for any protein target without the need to develop new reagents. Although the phage competition results are not a direct measure of the binding constant because the exact display level of each peptide on phage is not known, the phage ELISA method provides rapid access to the relative affinities of a large number of sequences.

Eight selectants from the F1 library selection and ten selectants from the F2 library selection were analyzed by competition phage ELISA, giving IC_{50} values ranging from low micromolar to submicromolar (Table 4). The three highest affinity sequences from the F1 library selection and the two most potent peptides from the F2 library selection were chosen for structure-function analysis as synthetic peptides. As a test of the methodology, we also synthesized two peptides based on sequences obtained in the naïve sort (Table 4).

Functional Analysis of Synthetic Peptides

Seven peptides were synthesized by standard Fmoc chemistry, corresponding to the second-generation phage-derived sequences and two parent sequences from the naïve selection (Table 4), with an additional glycine residue from the C-terminal linker to more precisely represent the peptide as displayed on phage. As synthetic peptides, the two families demonstrated distinct physical and functional characteristics. The peptides from Family 2 were poorly soluble; they underwent aggregation in aqueous solution and therefore yielded unreliable results. In contrast, the peptides from Family 1 were highly soluble; thus, they could be readily assayed and showed interesting activities (Table 5). This outcome highlights the fact that the properties of peptides displayed on phage can differ significantly from those of the same peptides free in solution. This obser-

Table 4. IC_{50} Values for the Interaction between IGF-1 and Phage-Displayed Peptides

Peptide[b]	Sequence	IC_{50} (μM)[a]
F1-Parent	DECLMSVEALKNCMG	–
F1-1[c]	RNCFESVAALRRCMY	0.1 ± 0.02
F1-2[c]	MDCLASVEALKWCMY	0.1 ± 0.07
F1-3[c]	FECLTSVEALRGCMY	0.3 ± 0.08
F1-4	FGCYESVAALRTCMY	0.3 ± 0.1
F1-5	HPCLESVGALKACMY	0.4 ± 0.06
F1-6	HDCFASVEALRRCMY	0.5 ± 0.13
F1-7	SDCFGSVEALKMCMY	0.7 ± 0.4
F1-8	SDCFESVEALRACMY	1.0 ± 0.7
F2-parent	ARCVVDAGGTWYCWA	–
F2-1[c]	ARCVVDAGGTWYCWA	0.5 ± 0.2
F2-2[c]	LGCASDLAGFWYCWA	0.3 ± 0.07
F2-3	WRCVDDLGGFQYCWA	1.0 ± 0.6
F2-4	ASCVADAGGGGYCWD	1.0 ± 0.7
F2-5	GGCASDLAGFRYCWE	1.0 ± 0.4
F2-6	VTCAADALGFLYCWE	2.0 ± 1.0
F2-7	LRCAEDLGGYFYCWA	1.4 ± 0.4
F2-8	YRCATDLAGFSYCWA	4 ± 5
F2-9	ARCARDAGGFYYCWA	1.0 ± 0.7
F2-10	WRCAQDAGGWTYCWA	1.0 ± 0.4

[a] The IC_{50} values are the concentrations of IGF-1 that blocked 50% of phage-displayed peptide binding to immobilized IGF-1 in a phage ELISA.
[b] The monovalent peptide-phage were from either library F1 or F2 (Table 3) and were selected for high-affinity binding to IGF-1.
[c] Sequences chosen for chemical synthesis and detailed functional analysis.

vation also emphasizes the importance of not performing excessive rounds of selection; if F2 family members had dominated in subsequent rounds of selection, then useful (i.e., soluble) peptide ligands for IGF-1 may not have been discovered.

The binding of IGF-F1-1 to IGF-1 was measured by using synthesized peptide to compete with IGF-F1-1 displayed on phage for binding to IGF-1 coated on a plate; the resulting IC_{50} was 7.2 ± 3.4 μM (data not shown). More functionally relevant assays were carried out with four IGF-1 binding partners to measure the antagonism of IGF-1 binding by the synthetic F1 peptides. For IGFBP-1 and IGFBP-3, we incubated each peptide with IGF-1 linked to P3 on phage and measured residual IGF-phage binding to each of the IGFBPs in a plate-based ELISA. Peptide IGF-F1-1, the highest affinity peptide when displayed on phage (Table 4), was also the most potent antagonist of each of the binding proteins (Table 5). Peptides IGF-F1-2 and IGF-F1-3 also blocked the binding of IGF-1 to IGFBP-1 and IGFBP-3, albeit with lower potency (see Table 5). The parent peptide, IGF-F1-parent, did not block the IGF-1 interaction with the IGFBPs within the detection limits of the experiment ($IC_{50} > 750$ μM).

A cell-based in vitro binding assay was used to test for peptide antagonism of the interaction between ^{125}I-IGF-I and the IGFR on cells in culture. The four peptides from family 1 were preincubated with 1 nM ^{125}I-IGF-1 prior to incubaton with MCF-7 cells that express IGFR. The cells were subsequently washed to remove unbound ^{125}I-IGF-1, and the residual radioactivity was measured after cell lysis to test the extent to which the peptides antagonized IGF-1 binding to IGFR. All F1 peptides, except for IGF-F1-parent, showed inhibition of ^{125}I-IGF-1 binding to IGFR on MCF-7 cells (Table 5). Finally, a cell-based in vitro kinase receptor activation (KIRA) assay was used to test the effects of the peptides on cultured cells transfected with functional insulin receptor (IR). In this assay, increased phosphorylation of receptors is seen after cells are stimulated with IGF-1, which binds and activates the IR with an EC_{50} of about 50 nM (data not shown). The results show that all the synthesized peptides inhibit receptor activation (Table 5). Interestingly, the IC_{50} values for the three matured peptides are equally potent across all four assays, whereas IGF-F1-parent shows weak activity only in the insulin receptor KIRA assay.

Structural Analysis of Synthetic Peptides

NMR chemical shift and $^3J_{HN\text{-}H\alpha}$ data for IGF-F1-1 deviated from random coil values (see Supplemental Material), suggesting that the peptide has a preferred conformation in solution. In particular, the $^3J_{HN\text{-}H\alpha}$ values for a number of residues (Ala8, Ala9, Leu10, Arg11, and Cys13) were less than 6.0 Hz, suggesting a helical conformation in this region of the peptide. The ensemble and minimized mean structure determined on the basis of NMR-derived experimental restraints are shown in Figure 2.

Analysis via the Kabsch and Sander secondary structure algorithm within PROCHECK [25] suggested that IGF-F1-1 contains an α helix from Val7 to Met14; Ser6 and Tyr15 are extensions of the main helix. Hydrogen bonding interactions consistent with the helical designation are observed in most of the structures within the ensemble. Cys3 and Phe4 adopt a reverse turn orientation (distorted Type I) in slightly more than half of the members of the ensemble. Residues Arg1, Asn2, and Gly16 are not well defined by the NMR restraints and

Table 5. IC_{50} Values for IGF-1 Binding Synthetic Peptides

		IC_{50} (μM)[a]			
Peptide	Sequence	IGFBP-1	IGFBP-3	IGFR	IR
F1-parent	DECLMSVEALKNCMGG-NH_2	>750	>750	>750	180
F1-1	RNCFESFVAALRRCMYG-NH_2	1.4	2.9	5.1	5.5
F1-2	MDCLASVEALKWCMYG-NH_2	16	9.2	15	33
F1-3	FECLTSVEALRGCMYG-NH_2	12	13	46	66

[a] The IC_{50} values are the concentrations of peptide that blocked 50% of the signal in assays designed to measure the interaction between IGF-1 and each of its ligands. The IGFBP values were from competitive ELISAs, the IGFR values were from MCF7 cell-based assays, and the IR values were from a KIRA assay (see Experimental Procedures for assay details).

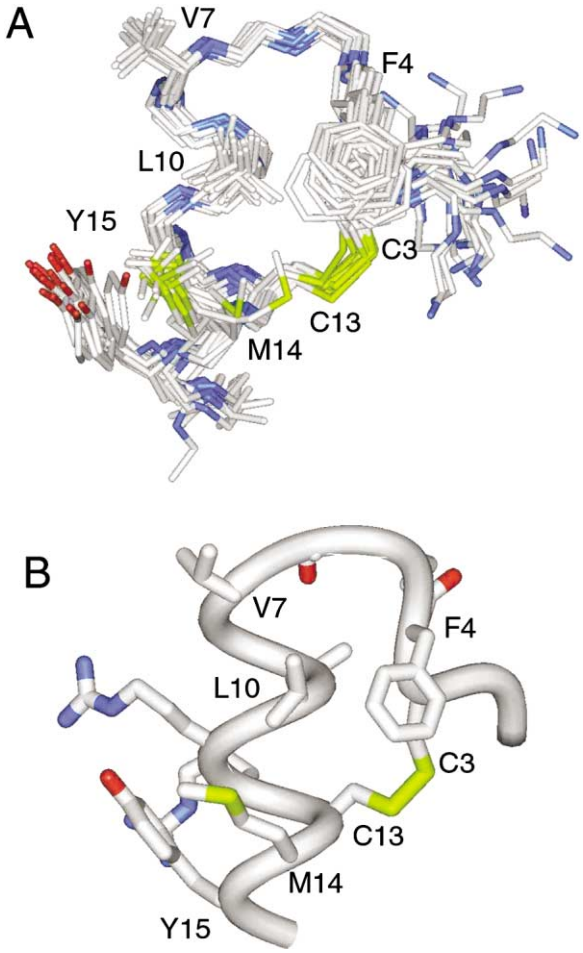

Figure 2. Three-Dimensional Structure of IGF-F1-1 in Solution

(A) Ensemble of 20 IGF-F1-1 structures superimposed by using the N, C^{α}, and C atoms of residues Phe4-Tyr15. The side chains of the hydrophobic residues on one face of the helix are shown.

(B) Minimized mean structure of IGF-F1-1 showing the heavy atoms of all side chains except Arg1 and Asn2.

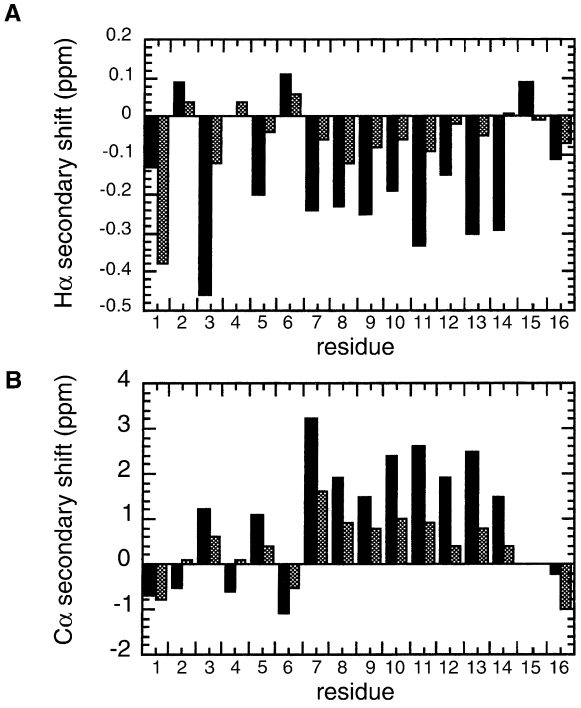

Figure 3. Comparison of Chemical Shifts for IGF-F1-1 and IGF-F1-Parent

(A) Secondary H^{α} chemical shifts (values for peptide − random-coil values) are plotted with solid and gray bars for IGF-F1-1 and IGF-F1-parent, respectively. Random-coil values were taken from [35].

(B) Secondary C^{α} chemical shifts (values for peptide − random-coil values) are plotted with solid and gray bars for IGF-F1-1 and IGF-F1-parent, respectively.

may be more flexible in solution than the other residues. There are extensive hydrophobic contacts between the side chains of Phe4, Val7, Leu10, Met14, and Tyr15. These residues also pack on top of the disulfide bond (residues Cys3 and Cys13). Presumably, the nonbonded interactions along this face of the helix help to stabilize the folded conformation of the peptide.

IGF-F1-parent, the precursor to IGF-F1-1, was also studied by NMR (Supplemental Material). In the region of the helix observed in IGF-F1-1, the H^{α} and C^{α} chemical shifts of IGF-F1-parent are systematically shifted from random coil values in a direction consistent with a helix also being present. However, the shifts are not as extreme as in IGF-F1-1 (Figure 3). The values of $^{3}J_{HN\text{-}H\alpha}$ in the helix are also smaller in IGF-F1-parent (mean for Val7-Met14 = 6.4 Hz) than in IGF-F1-1 (mean for Val7-Met14 = 5.5 Hz; see Supplementary Material). In addition, very few of the medium-range ROEs present in IGF-F1-1 are observed in the parent peptide. These data all suggest that although IGF-F1-parent may adopt a structure similar to that of IGF-F1-1, this conformation may be less highly populated and therefore less well defined and more dynamic in nature. Given this, a structure was not determined for IGF-F1-parent.

Discussion

In the present work, we have described a modified protocol for the identification of novel peptide ligands by using phage-displayed libraries. The aim of these modifications was to decrease the time and expense of identifying ligands and also to identify a number of classes of potential ligands for further optimization and characterization. A key component was the use of generic (i.e., target-independent) procedures and analytical tools that are fast and highly automatable; such tools included plate-based ELISA readouts and DNA sequencing. Initial searching was performed with a single large multiple-scaffold library by the use of polyvalent display to boost target binding via avidity. Single-point ELISA tests were used to assess the binding competency of clones and to prevent loss of diversity through over-sorting. Once the majority of the clones bound to target, a relatively large number of binding clones were sequenced, and these data were used to design focused second-generation libraries that could be exhaustively sampled for the highest affinity ligands. The use of monovalent display was critical for this step to allow more stringent selection

of ligands whose affinity could be measured directly by phage ELISA, unimpeded by avidity effects. Expensive and time-consuming peptide synthesis was not performed until after the final affinity maturation step.

Using this optimized protocol, we sought to identify ligands to the growth factor IGF-1. The initial naïve selection yielded peptides from six distinct scaffolds, with one scaffold, CX_9C, dominating the sequenced clones. Of the two families of related CX_9C peptides derived from the sequences, the family F1 contained a more highly conserved motif (Table 3). Sequence alignments were made between the five F1 and F2 peptides that were synthesized and all human IGFBPs, IGF-R1, IGF-R2, and IR. In no case was an identity of more than 30% observed, and in nearly all cases gaps were required to obtain any alignment identity greater than 15% (our unpublished data). Thus, the selection sequences are novel epitopes for binding IGF-1 and are not simply linear fragments of known IGF binding proteins.

Synthetic peptides derived from affinity maturation of F1 bind to IGF-1 with good potency and have a well-defined solution structure (Table 5 and Figure 2). The other family, F2, is based on a less well-defined consensus sequence (Table 2); synthetic peptides from this family had a tendency to aggregate or have limited solubility, making characterization of their structure or IGF-1 affinity problematic. The affinity-matured F2 sequences that could be assessed by phage ELISA were found to be less potent than the F1 sequences (Table 4). The lower level of consensus in naïve F2 sequences suggests that we might have started with a sub-optimal binding epitope that could not evolve into a useful high-affinity ligand. The generality of the correlation between high consensus in naïve sorts and higher potency in the final optimized affinity remains to be seen but will become clearer as we gain more experience with using this modified procedure for ligand discovery.

Although selection was based only on the ability of peptide-phage to bind to IGF-1, the sequence of IGF-F1-1 also folds into a stable, compact structure. We are aware of two examples [7, 26] in which the conformation observed for a phage-derived peptide in solution is similar to that present when the peptide is bound to the target protein. Selection of a peptide that has a stable fold in solution and does not change significantly upon binding to its target provides an energetic benefit to binding because the association will not lead to a loss of conformational entropy. Accordingly, structured peptides derived from phage [7, 9, 26–28] retain their native conformations when they are bound to the target protein, suggesting that the structure of IGF-F1-1 bound to IGF-1 is preserved. Interestingly, the fold observed in IGF-F1-1 (N-terminal loop, C-terminal helix) has been observed in a number of other phage-selected protein ligands, suggesting an inherent stability of this motif [7, 9, 26–28]. Moreover hydrophobic residues on one face of the helix provide important contacts for binding to the target protein in several of these examples [7, 26–28]. On the basis of these prior findings, we hypothesize that the hydrophobic patch depicted in Figure 2 probably represents the surface that IGF-F1-1 uses to bind to IGF-I. An important observation is that the hydrophobic patch is formed by the consensus residues (Phe4, Val7, Leu10, and Met14) identified in the initial selection (Table 2). Although these residues are identical or of similar character in the naïve selected peptide IGF-F1-parent, the affinity for IGF-1 is at least an order of magnitude less than that of the affinity-matured peptide IGF-F1-1, as judged from the IGFBP1 and IGFBP3 competition binding studies (Table 5). The working model we have developed for the evolution process is that approximate binding solutions are found within incomplete, but highly diverse, polyvalent libraries that allow us to identify binding motifs and, therefore, to design directed libraries with complete coverage of relevant sequence space. The directed libraries contain peptides with the binding epitope displayed on a more highly structured scaffold, so that these peptides bind IGF-1 with higher affinity and are selected preferentially. Although this model agrees with our results for IGF-1 ligands, identification and optimization of ligands to other targets will be required to test if it is generally applicable. If our model is correct, the discovery of a tight binding peptide by this methodology requires the presence of a highly conserved binding core in the original selection. The probability of this occurring is dependent on the properties of target, choice of libraries, and completeness with which the libraries are sampled. In this case, by using the maximum diversity available to us, we found consensus motifs that led to potent binders by sequencing fewer than 100 clones.

Previous mutagenesis studies have identified distinct epitopes on IGF-1 that interact with binding proteins and receptors [20, 21]. These data suggest that the IGFBPs and receptors bind to different and distinct regions of the growth factor. However, binding to IGFBP-1 or other binding proteins does inhibit receptor signaling by IGF-1 [17]. This implies that IGFBP binding partially blocks the receptor binding site or induces a conformational change in IGF-1 that influences the display of the receptor epitope; the latter seems reasonable given the highly dynamic nature of IGF-1 in solution [29]. Interestingly, IGF-F1-1 antagonizes IGF-1 binding to IGFBPs and cell-surface receptors with similar potency. As with the mutagenesis results, this result may be rationalized in two ways (direct steric blocking or an allosteric perturbation of the binding site), and both may be occurring. Structural characterization of IGF-1 in complex with IGF-F1-1 is currently underway. These experiments will reveal the molecular details of how the phage-derived peptide recognizes IGF-1. Moreover, by making comparisons to previous structures of IGF-1 [23, 24], and to mutagenesis data [20–22], we hope to shed light on how IGF-1 activity is controlled by binding to peptides or binding proteins.

Significance

In this report we described methodology that maximizes the diversity that can be sampled by the use of peptide-phage libraries to rapidly identify and assess peptide ligands for any target. The major points are (1) sorting a large ($>10^{11}$) library of naïve peptides for binding to a target protein, as well as sequencing a moderate number of clones, can yield a recognizable

motif that is responsible for binding; (2) the naïve polyvalent sort identified the residues necessary for forming the binding surface, and the subsequent evolution of directed sub-libraries enabled us to access all the sequences that present this motif; (3) peptides that presented a more structured binding interface were selected in the second-generation monovalent sort; (4) the increase in potency is probably due to preorganization of the binding interface; (5) the most potent peptide derived from monovalent selection, IGF-F1-1, antagonizes the interaction of IGF-1 with the IGFBPs and cell-surface receptors with low micromolar IC_{50} values; (6) all three peptides tested from the monovalent sort showed similar activity profiles against all of the IGF-1 ligands tested; and (7) sequences derived from the naïve sort could show different activities and selectivities and suggest additional binding motifs.

Experimental Procedures

Construction of Polyvalent Naïve Peptide-Phage Libraries

Libraries were constructed by previously described methods [5], with a phagemid vector containing an IPTG-inducible P_{tac} promoter driving the expression of open-reading frames encoding fusion proteins of the following form: the StII secretion signal (MKKNIAFLLAS MFVFSIATNAYA), followed by a random peptide, followed by a linker (GGGSGGG), followed by the M13 gene-8 major coat protein (AEG DDPAKAAFNSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSKAS). Twenty-two different peptide libraries were constructed (Table 1), and peptide-phage displaying the libraries were purified by precipitation with PEG/NaCl, as described previously [5], and stored at −70°C.

Selection of IGF-1 Binding Peptides from Naïve Peptide-Phage Libraries

IGF-1 was obtained as described previously [24]. Immunosorbant plates (Nunc Maxisorp) were coated with 5 μg/ml of IGF-I in 50 mM sodium carbonate buffer (pH 9.6) for 1 hr at room temperature, followed by blocking for 1 hr with 0.2% BSA in phosphate-buffered saline (PBS). The plates were washed with PBS and 0.05% Tween 20.

Phage from 22 naïve peptide-phage libraries (Table 1) were pooled and cycled through three rounds of binding selection on immunosorbant plates coated with IGF-1. In the first round, 4.8 ml of phage cocktail (~10^{13} phage/ml in PBS, 0.05% Tween 20, and 0.2% BSA [BSA/Tween buffer]) were added to 48 coated wells (100 μl/well). After 2 hr incubation at room temperature with shaking, the plate was washed with PBS and 0.05% Tween 20 to remove unbound phage. Bound phage were eluted with 0.2 M glycine (pH 2.0) (100 μl/well), and the phage eluant was neutralized by the addition of 1/6 volume of 1.0 M Tris (pH 8.0). The eluted phage were amplified overnight by propagation in *E. coli* XL1-blue cells (Stratagene) with M13-VCS helper phage (Stratagene) and harvested by precipitation with PEG/NaCl as described previously [5]. The selection procedure for round three was identical to that for round one, whereas round two differed only in the use of 0.2% casein in place of BSA in both the blocking buffer and the phage cocktail. Individual phage clones at each round were analyzed for specific binding to IGF-1 with single-point phage ELISAs (see below). Positive clones, those which bound to IGF-1 but not BSA, were subjected to DNA sequence analysis (see below).

Second-Generation Affinity Maturation of IGF-1 Binding Peptides

Second generation libraries were constructed by previously described methods [5], with a phagemid (phGHam-g3) [4] containing an alkaline phosphotase promoter (PphoA) driving the expression of fusion proteins of the following form: StII secretion signal, followed by the second-generation peptide library, followed by the linker sequence, followed by human growth hormone, followed by the C-terminal domain of the M13 gene-3 minor coat protein. Four second-generation peptide libraries were constructed as shown in Table 3.

Phage from the second-generation libraries were cycled through four rounds of binding selections as described above. For rounds one and three, immobilized IGF-1 was used as the capture target, with 0.2% BSA in the blocking buffer and the phage cocktail. Rounds two and four used anti-hGH monoclonal antibody 3F6.B1.4B1 [5] as the capture target, with 0.2% casein in the blocking buffer and the phage cocktail. The clones were assayed for binding to IGF-F1-1, BSA, and anti-hGH monoclonal antibody 3F6.B1.4B1. A fifth round of selection was conducted in which the harvested phage pool was incubated with 1.0 nM biotinylated IGF-1 (bio-IGF) in BSA/Tween buffer for 2 hr. The phage/bio-IGF solution was added to streptavidin-linked magnetic beads (Promega) (previously blocked with BSA). After 30 min, the magnetic beads were washed with PBS and 0.05% Tween 20 to remove unbound phage. The remaining phage were eluted with 0.1 M HCl and neutralized by the addition of 1/6 volume of 1.0 M Tris (pH 8.0). Individual clones from round 5 were analyzed by single-point phage ELISAs (see below). Positive monovalent phage clones were identified as those that bound to both IGF-1 and anti-hGH antibody but not to BSA. Positive clones were subjected to DNA sequence analysis (see below).

Single-Point Phage ELISA

To facilitate the rapid analysis of hundreds of unique phage clones, we used a 96-well format for phage growth and ELISAs. Individual *E. coli* XL1-blue colonies harboring phagemids were picked into 500 μl of growth medium (2YT supplemented with 50 μg/ml carbenicillin and 10^{10} pfu/ml VCS-M13 helper phage) in 1.2 ml culture tubes in a 96-well format. After overnight growth, the tube racks were centrifuged for 10 min at 2.5k rpm in a Sigma 3–12 centrifuge equipped with an Nr.11222 rotor. The phage supernatants (300 μl) were transferred to 96-well plates and used directly in single-point phage ELISAs.

For phage ELISAs, 96-well maxisorp immunoplates (NUNC) were coated with 100 μl of 5 μg/ml capture target protein, blocked with BSA, and washed (as described above). Phage supernatants (50 μl) were added to individual wells and incubated for 1 hr with shaking at room temperature. The plates were washed eight times with PBS and 0.05% Tween 20, incubated with 50 μl of 1:10,000 horse radish peroxidase/anti-M13 antibody conjugate (Pharmacia) in BSA/Tween buffer for 30 min, and then washed eight times with PBS and 0.05% Tween 20 and two times with PBS. Plates were developed with a tetramethylbenzidine substrate (TMB, Kirkegaard and Perry, Gaithersburg, MD), quenched with 1.0 M H_3PO_4 (50 μl), and read spectrophotometrically at 450 nm.

DNA Sequencing

DNA sequencing was performed in a 96-well format. Culture supernatants containing phage particles (see above) were heat sterilized at 60°C for 1 hr and then used as the template (2 μl) for a PCR that amplified a 300 base pair fragment surrounding the peptide library sequence. Standard PCR conditions were used, with Taq polymerase (1.25 units in a 25 μl reaction, Applied Biosystems) as the amplifying enzyme. Bipartite PCR primers were used to simultaneously amplify the DNA and to add M13(–21) and M13R universal sequencing primers to the 5′ and 3′ ends of the fragment, respectively. To degrade excess primers and nucleotides, we treated the PCR (5 μl) with shrimp alkaline phosphatase (1 unit) and exonuclease-1 (10 units) (USB) for 15 min at 37°C and 15 min at 80°C. The treated PCR was used as the template in Big-Dye terminator sequencing reactions (with either the M13(–21) or M13R universal sequencing primer). The sequencing reactions were analyzed on an ABI Prism 3700 DNA analyzer equipped with a 96-capillary array.

Affinity Measurement by Competitive Phage ELISA

A modified phage ELISA was used to estimate the binding affinities of selected second-generation IGF-1 binding peptides from above. Phage ELISAs were carried out on plates coated with IGF-1, as described above. Peptide-displaying phage were serially diluted, and binding was measured to determine a phage concentration giving ~50% of the signal at saturation.

A fixed, subsaturating concentration of peptide-phage was mixed

with serial dilutions of IGF-1, incubated for 1.0 hr, and then transferred to assay plates coated with IGF-1. After 30 min of incubation, the plates were washed and developed, as described above. The binding affinities of the peptides for IGF-1 were determined as IC_{50} values (Kaleidagraph) defined as the concentration of IGF-1 that blocked 50% of the peptide-phage binding to the immobilized IGF-1. An ELISA format was also used to measure the binding of synthetic peptide to IGF-1. Phage displaying IGF-F1-1 on P3 (no hGH) were bound to a plate coated with IGF-1. Serial dilutions of synthetic peptide were used to determine the concentration at which 50% of the phage were displaced.

Peptide Synthesis

Peptides were synthesized by either manual or automated (Milligen 9050) solid-phase synthesis at 0.2 mM scale on PEG-polystirene resin via Fmoc chemistry. Purification was as previously described [26].

Inhibition of IGF-I Binding to IGFBPs

E. coli XL1-Blue cells harboring the phagemid pIGF-g3 were used to produce phage particles displaying human IGF-1, as described previously [22]. For each IGF-1 binding synthetic peptide, competitive phage ELISAs were performed to determine IC_{50} values, defined as the concentration of synthetic peptide that blocked 50% of the IGF-1-phage binding to immobilized IGFBP-1 or IGFBP-3. The procedures were identical to those in section 5.6, except that plates were coated with IGFBP-1 or IGFBP-3, and subsaturating concentrations of IGF-1-phage were preincubated with serial dilutions of synthetic peptides.

Inhibition of IGF-1-Induced Activation of the Insulin Receptor

A kinase receptor activation (KIRA) assay was developed to measure the phosphorylation of the human insulin receptor (hIR) in response to exogenous IGF-1. The methods were modified from those of a previously described KIRA assay [30]. Chinese hamster ovary cells transfected with hIR (TRY-IR 5.3 cells) were grown overnight at 37°C in 96-well plates with PS/20 medium. Supernatants were decanted, and stimulation medium (PS/20, 0.5% BSA, and 25 nM IGF-1) was added. After 15 min at 37°C, the stimulation medium was removed, and cells were lysed by the addition of lysis buffer (50 mM HEPES, 150 mM NaCl, 0.5% Triton X-100, 1.0 mM 4-[2-aminoethyl]benzenesulfonyl fluoride hydrochloride, 30 kIU/ml aprotinin, 50 uM leupeptin, and 2.0 mM sodium orthovanadate).

The lysates were transferred to immunosorbant plates that had been coated overnight with 100 μl anti-hIR antibody (2 μg/ml) and blocked with BSA. After 2 hr incubation to allow for capture of solubilized hIR from the lysates, the plates were washed and anti-phosphotyrosine antibody 4G10 (Upstate Biologicals Inc.) was added to detect activated hIR. After 2 hr incubation, the plates were developed with streptavidin conjugated to HRP and TMB substrate as described above.

To measure the inhibitory effects of IGF-1 binding synthetic peptides on hIR activation, the stimulation medium was supplemented with serial dilutions of peptide stocks prior to the addition of the medium to the cell cultures. The IC_{50} was defined as the concentration of peptide that blocked 50% of the hIR phosphorylation.

Displacement of IGF-I on MCF7 Cells

MCF7 (ATCC-HTB, Bethesda, MD) breast carcinoma cells, which express IGF receptors and insulin receptors [31], were used to detect the binding of IGF-I to cell-surface receptors. Cells were passaged weekly in media containing a 10 mM HEPES (pH 7.2) 1:1 mixture of high glucose DMEM/Ham's F12, with 10% fetal bovine serum. Cells were plated and grown overnight at 37°C and 5% CO_2. Serial dilutions of synthetic peptides were preincubated with 2.0 nM ^{125}I-labeled IGF-I for 40 min at 37°C. The samples were then added to MCF7 cells for 30 min. The cells were washed with media and lysed with 1.0 M NaOH. Radioactivity was counted for 15 min to quantitate cell-bound ^{125}I-labeled IGF-1. At these assay concentrations, IGF-1 (2.0 nM) binds to IGF receptors but does not bind significantly to insulin receptors (data not shown). The IC_{50} was defined as the concentration of peptide that displaces 50% of the ^{125}I-labeled IGF-1 from the MCF7 cells.

Structure Determination of IGF-F1-1 by NMR

^{1}H NMR data were collected in H_2O (with 10% v/v ^2H_2O) or ^2H_2O solution at 30°C (pH 5.1) and 5.0 mM concentration. In addition to one-dimensional spectra, two-dimensional double-quantum-filtered correlation spectra (2QF-COSY), total correlation spectra (TOCSY; t_m = 90 ms), and rotating-frame Overhauser effect spectra (ROESY; t_m = 150 ms) were collected [9, 32]. After lyophilization and dissolution of the peptide in ^2H_2O, a 2D ROESY spectrum and a COSY spectrum with a 35° mixing pulse [32] were acquired. Complete ^{1}H resonance assignments were obtained from these data by standard methods [33] and are listed in the Supplemental Material. Natural abundance ^{13}C HMQC and ^{13}C-HMQC-TOCSY spectra were also collected to facilitate assignment of the ^{13}C resonances [32]. Distance restraints were generated from H_2O and D_2O ROESY spectra, and dihedral angle restraints were derived from $^3J_{HN\text{-}H\alpha}$ or $^3J_{H\alpha\text{-}H\beta}$ measured in H_2O 2QF-COSY or D_2O COSY-35 spectra, respectively, processed to high digital resolution in F_2 [32]. Additional ϕ and ψ restraints were generated from the H^α, C^α, and C^β chemical shifts with the program TALOS [34]. Restraints were set to the value determined from TALOS ± the larger of either twice the TALOS standard deviation or 40°. Hydrogen bond restraints were not employed at any stage of the calculation process. The ROESY coupling constant and chemical shift data were used to generate 91 distance and 26 dihedral angle restraints that were used as input to structure calculations. One hundred initial structures were calculated with the hybrid distance geometry/simulated annealing program DGII [35]; 80 of these were further refined by restrained molecular dynamics with the AMBER all-atom force field implemented in DISCOVER as described previously [36]. Twenty conformations of lowest restraint violation energy were chosen to represent the solution conformation. Mean coordinates generated from this ensemble were energy minimized under the influence of the experimental restraints to produce a minimized mean structure. Restraint numbers and structural statistics for the ensemble and minimized mean structure are presented in the Supplemental Material. Random-coil ^{13}C chemical shifts for IGF-F1-1 and IGF-F1-parent were determined from values tabulated previously [37]. The resulting ensemble of 20 structures converged to a single global fold (mean rmsd from the mean structure of 0.31 ± 0.06 Å and 1.18 ± 0.15 Å for backbone or all heavy atoms, respectively, of residues 3–15). The structures agreed with the input data very well (no distance restraint violations greater than 0.07 Å and no dihedral angle violations greater than 1.0°), and had good covalent geometry as judged by the program PROCHECK [25].

Supplementary Material

NMR resonance assignments for peptides IGF-F1-1 and IGF-1-parent are available in Supplementary Tables 1 and 2. Supplementary Table 2 contains details of the input restraints used to calculate the structure of IGF-F1-1 and quality and precision statistics for the resulting ensemble.

Acknowledgments

The authors thank the following: Daniel Tran for iodination of IGF-1; the DNA synthesis and sequencing groups at Genentech for preparation and analysis of DNA; Cliff Quan, Jeff Tom, and Anna Fedorova for peptide synthesis; Tom Roth for assistance with the phage work; and Henry B. Lowman and Yves Dubaquie for many helpful discussions.

Received: October 24, 2001
Revised: January 31, 2002
Accepted: February 12, 2002

References

1. Downward, J. (2001). The ins and outs of signaling. Nature *411*, 759–762.
2. Sidhu, S.S. (2000). Phage display in pharmaceutical biotechnology. Curr. Opin. Biotechnol. *11*, 613–616.
3. Cochran, A.G. (2000). Antagonists of protein-protein interactions. Chem. Biol. *9*, R85–R94.
4. Lowman, H.B., and Wells, J.A. (1993). Affinity maturation of

human growth hormone by monovalent phage display. J. Mol. Biol. *234*, 564–578.
5. Sidhu, S.S., Lowman, H.B., Cunningham, B.C., and Wells, J.A. (2000). Phage display for selection of novel binding peptides. Methods Enzymol. *328*, 333–363.
6. Olivera, B.M., Hillyard, D.R., Marsh, M., and Yoshikami, D. (1995). Combinatorial peptide libraries in drug design: lessons from venomous cone snails. Trends Biotechnol. *13*, 422–426.
7. Dennis, M.S., Eigenbrot, C., Skelton, N.J., Ultsch, M.H., Santell, L., Dwyer, M.A., O'Connell, M.P., and Lazarus, R.A. (2000). Peptide exosite inhibitors of factor VIIa as anticoagulants. Nature *404*, 465–470.
8. Fairbrother, W.J., Christinger, H.W., Cochran, A.G., Fuh, G., Keenan, C.J., Quan, C., Shriver, S.K., Tom, J.Y.K., Wells, J.A., and Cunningham, B.C. (1998). Novel peptides selected to bind vascular endothelial growth factor target the receptor-binding site. Biochemistry *37*, 17754–17764.
9. Lowman, H.B., Chen, Y.M., Skelton, N.J., Mortensen, D.L., Tomlinson, E.E., Sadick, M.D., Robinson, I.C.A.F., and Clark, R.G. (1998). Molecular mimics of insulin-like growth factor 1 (IGF-1) for inhibiting IGF-1: IGF-binding protein interactions. Biochemistry *37*, 8870–8878.
10. Wrighton, N.C., Farrell, F.X., Chang, R., Kashyap, A.K., Barbone, F.P., Mulcahy, L.S., Johnson, D.L., Barrett, R.W., Loliffe, L.K., and Dower, W.J. (1996). Small peptides as potent mimetics of the protein hormone erythropoietin. Science *273*, 458–463.
11. Daughaday, W.H., and Rotwein, P. (1989). Insulin-like growth factors I and II. Peptide, messenger ribonucleic acid and gene structures, serum, and tissue concentrations. Endocr. Rev. *10*, 68–91.
12. Liu, J.L., Grinberg, A., Westphal, H., Sauer, B., Accili, D., Karas, M., and LeRoith, D. (1998). Insulin-like growth factor-I affects perinatal lethality and postnatal development in a gene dosage-dependent manner: manipulation using the Cre/loxP system in transgenic mice. Mol. Endocrinol. *12*, 1452–1462.
13. Sjogren, K., Liu, J.L., Blad, K., Skrtic, S., Vidal, O., Wallenius, V., LeRoith, D., Tornell, J., Isaksson, O.G., Jansson, J.O., et al. (1999). Liver-derived insulin-like growth factor 1 (IGF-1) is the principal source of IGF-1 in blood but is not required for postnatal body growth in mice. Proc. Natl. Acad. Sci. USA *96*, 7088–7092.
14. Schlechter, N.L., Russell, S.M., Spencer, E.M., and Nicoll, C.S. (1986). Evidence suggesting that the direct growth-promoting effect of growth hormone on cartilage in vivo is mediated by local production of somatomedin. Proc. Natl. Acad. Sci. USA *83*, 7932–7934.
15. Isaksson, O.G.P., Jansson, J.-O., and Gause, I.A.M. (1982). Growth hormone stimulates longitudinal bone growth directly. Science *216*, 1237–1239.
16. Jones, J.I., and Clemmons, D.R. (1995). Insulin-like growth factors and their binding proteins: biological actions. Endocr. Rev. *16*, 3–34.
17. Chan, J.M., Stampfer, M.J., Giovannucci, E., Gann, P.H., Jing, M., Wilkinson, P., Hennekens, C.H., and Pollak, M. (1998). Plasma insulin-like growth factor-1 and prostate cancer risk: a prospective study. Science *279*, 563–566.
18. Arteaga, C.L. (1992). Interference of the IGF system as a strategy to inhibit breast cancer growth. Breast Cancer Res. Treat. *22*, 101–106.
19. Loddick, S.A., Liu, X.-J., Lu, Z.-X., Liu, C., Behan, D.P., Chalmers, D.C., Foster, A.C., Vale, W.W., Ling, N., and De Souza, E.B. (1998). Displacement of insulin-like growth factors from their binding proteins as a potential treatment for stroke. Proc. Natl. Acad. Sci. USA *873*, 1894–1898.
20. Bayne, M.L., Applebaum, J., Underwood, D., Chicchi, G.G., Green, B.G., Hayes, N.S., and Cascieri, M.A. (1988). Structural analogs of human insulin-like growth factor 1 with reduced affinity for serum binding proteins and the type 2 insulin-like growth factor receptor. J. Biol. Chem. *264*, 11004–11008.
21. Bayne, M.L., Applebaum, J., Miller, R.E., Chicchi, G.G., Green, B.G., Hayes, N.S., and Cascieri, M.A. (1990). The roles of tyrosines *24*, 31, and 60 in the high affinity binding of insulin-like growth factor-1 to the type 1 insulin-like growth factor receptor. J. Biol. Chem. *265*, 15648–15652.
22. Dubaquié, Y., and Lowman, H.B. (1999). Total alanine-scanning mutagenesis of insulin-like growth factor 1 (IGF-1) identifies differential binding epitopes for IGFBP-1 and IGFBP-3. Biochemistry *38*, 6386–6396.
23. Zeslawski, W., Beisel, H.-G., Kamionka, M., Kalus, W., Engh, R.A., Huber, R., Lang, K., and Holak, T. (2001). The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5. EMBO J. *20*, 3638–3644.
24. Vadjos, F.F., Schaffer, M.L., Deshayes, K., Liu, J., Skelton, N.J., and de Vos, A.M. (2001). Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions. Biochemistry *40*, 11022–11029.
25. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
26. Skelton, N.J., Chen, Y.M., Mortensen, D.L., Dubree, N., Quan, C., Jackson, D.Y., Cochran, A., Zobel, K., Deshayes, K., Baca, M., et al. (2001). Structure-function analysis of a phage display-derived peptide that binds to insulin-like growth factor binding protein 1. Biochemistry *40*, 8487–8498.
27. Pan, B., Russell, S., Li, B., Tom, J.Y.K., Cochran, A.G., and Fairbrother, W.J. (2002). The solution structure of a phage-derived peptide antagonist in complex with vascular endothelial growth factor. J. Mol. Biol. *316*, 769–787.
28. Eckert, D.M., Malashkevich, V.N., Hong, L.H., Carr, P.A., and Kim, P.S. (1999). Inhibiting HIV-1 entry: Discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell *99*, 103–115.
29. Cooke, R.M., Harvey, T.S., and Campbel, I.D. (1991). I Solution structure of human-like growthfactor 1: a nuclear magnetic resonance and restrained molecular dynamics study. Biochemistry *30*, 5484–5491.
30. Sadick, M.D., Quarmby, A.I.V., McCoy, A., Canova-Davis, E., and Ling, V. (1999). Kinase receptor activation (KIRA): a rapid and accurate alternative to end-point bioassays. J. Pharm. Biomed. Anal. *19*, 883–891.
31. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Handbook. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
32. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., and Skelton, N.J. (1995). Protein NMR Spectroscopy: Principles and Practice. (New York: Academic Press).
33. Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids. (New York: Wiley).
34. Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone dihedral angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR *13*, 289–302.
35. Havel, T.F. (1991). An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. Prog. Biophys. Mol. Biol. *56*, 43–78.
36. Skelton, N.J., Garcia, K.C., Goeddel, D.V., Quan, C., and Burnier, J.P. (1994). Determination of the solution structure of guanylin. Biochemistry *3*, 13581–13592.
37. Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., and Sykes, B.D. (1995). ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of common amino acids: investigation of nearest neighbor effects. J. Biomol. NMR *5*, 67–81.

Accession Numbers

The ensemble of 20 structures representing the solution conformation of IGF-F1-1 have been deposited with the protein data bank (accession number 1LB7).