Abstract

Miniature Protein Ligands for CBP KIX:
Selection of Transcriptional Activation Domains

Heather Marie Volkman

2005

This dissertation describes the selection and characterization of miniature protein ligands for the protein surface of transcriptional coactivator CBP. Ligands were identified using protein grafting in combination with phage display where the CBP-binding residues from transcription factor CREB were displayed on the stabilized helical scaffold aPP. Chapter 1 illustrates the ability to identify miniature protein ligands for protein surfaces with the incorporation of a post-translational modification step in the selection process. This selection identified phosphorylated and unphosphorylated peptide ligands that bind to CBP KIX with high affinity and specificity. The binding affinity of the best phosphorylated (PPKID4\textsuperscript{P}) and unphosphorylated (PPKID6\textsuperscript{U}) ligands were explored in greater detail using a panel of CBP KIX variants where residues within and around the binding cleft were individually substituted with alanine. The results support a model in which PPKID4\textsuperscript{P} binds CBP KIX in a manner that closely resembles that of the natural ligand CREB KID, but that PPKID6\textsuperscript{U} binds an overlapping, yet distinct region of the protein. Additionally, the binding affinity of a set of PPKID4\textsuperscript{P} variants for CBP KIX was measured to illustrate the possibility that PPKID4\textsuperscript{P} folds into an aPP-like conformation upon binding to CBP KIX. Chapter 2 investigates the potential of PPKID ligands to act as transcriptional activation domains when fused to a heterologous DNA-binding domain. Transcriptional activation assays performed in HEK293 cells using
GAL4 DNA-binding domain fusion proteins of PPKID4 and PPKID6 indicate that PPKID4\textsuperscript{p} functions as a potent activator of p300/CBP-dependent transcription. PPKID6\textsuperscript{p} is a less potent transcriptional activator in this context than is PPKID4\textsuperscript{p} despite the similarity of their affinities for CBP KIX. This final result suggests that thermodynamic binding affinity is an important, although not exclusive, criterion controlling the level of CBP-dependent transcriptional activation.
Miniature Protein Ligands for CBP KIX:
Selection of Transcriptional Activation Domains

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Heather Marie Volkman

Dissertation Director: Alanna Schepartz

May 2005
# Table of Contents

Abstract ................................................................................................................................. i

Title Page ............................................................................................................................. iii

Copyright ............................................................................................................................... iv

Table of Contents .................................................................................................................. v

List of Figures ....................................................................................................................... xi

List of Tables ......................................................................................................................... xiv
Chapter 1 Selection of Miniature Protein Ligands for CBP KIX

Introduction ......................................................................................................................... 2

Protein-protein interactions .............................................................................................. 2

Protein-protein interfaces ................................................................................................. 2

Protein surface recognition .............................................................................................. 5

Protein Minimization ........................................................................................................ 5

Peptide Scaffolds .............................................................................................................. 7

Peptoid scaffolds ............................................................................................................... 9

β-peptide scaffolds ........................................................................................................... 10

Macrocyclic scaffold ......................................................................................................... 10

Therapeutic antibodies ..................................................................................................... 13

Small molecule inhibitors .................................................................................................. 13

Alpha helix stabilization .................................................................................................... 15

Disulfide and other covalent bonds .................................................................................. 15

Non-peptidic scaffolds ....................................................................................................... 16

Avian Pancreatic Polypeptide (aPP) ................................................................................ 17

Protein Grafting ................................................................................................................ 20

Using the α-helix of aPP to target DNA .......................................................................... 20

Using the α-helix of aPP to target proteins ..................................................................... 21

Using the PPII helix of aPP to target proteins ................................................................. 22

Extension of protein grafting to flat protein surfaces ......................................................... 22

The transcription factor CREB ......................................................................................... 23

KID-KIX interaction .......................................................................................................... 23
Discussion ........................................................................................................................................ 122

Phosphorylated miniature protein ligands for CBP KIX ....................................................... 123

Comparison to unphosphorylated counterparts ............................................................... 125

Unphosphorylated miniature protein ligands for CBP KIX .................................................. 126

PPKID specificity ................................................................................................................ 128

Structure of PPKID ............................................................................................................. 129

Binding mode and orientation of CBP ligands: PPKID4” .................................................. 131

Binding mode and orientation of CBP ligands: PPKID6” .................................................. 135

PPKID4” Folding ................................................................................................................ 136

Experimental ...................................................................................................................... 137

General ................................................................................................................................. 137

Materials .............................................................................................................................. 137

Media and Buffers ............................................................................................................... 138

Expression and Purification of GST-KIX and KIX variants ............................................. 139

Phage Display Selections .................................................................................................... 140

Peptide Sequences for Control Peptides ............................................................................ 142

Peptide Synthesis, Purification and Modification .............................................................. 142

Fluorescence Polarization .................................................................................................... 143

Single sample format .......................................................................................................... 143

Multi-well format ................................................................................................................ 144

Fluorescence Polarization Competition .............................................................................. 145

Circular Dichroism .............................................................................................................. 146

References ............................................................................................................................. 147
Chapter 2 Transcriptional Activation Potential of PPKID Ligand

Introduction.......................................................................................................................... 153

Regulation of transcription................................................................................................. 153

Mechanism of activation..................................................................................................... 154

Artificial transcription factors .......................................................................................... 155

Natural DNA-binding domains ......................................................................................... 156

Artificial DNA-binding domains....................................................................................... 156

Natural activation domains ............................................................................................... 162

Artificial activation domains............................................................................................ 163

Correlation between affinity and activation....................................................................... 164

CBP as a transcriptional coactivator .................................................................................. 165

Miniature protein ligands for the KIX domain of CBP ................................................... 166

Results.................................................................................................................................. 167

Experimental goals............................................................................................................. 167

Constructs for transcription reporter assay ................................................................. 168

Expression of constructs in HEK293 cells ...................................................................... 173

Phosphorylation of constructs in HEK293 cells ............................................................. 173

Transcriptional activation by PPKID ............................................................................... 178

Activation pathway of PPKID.......................................................................................... 181

Effects of oncoprotein E1A .............................................................................................. 181

Effects of transcriptional coactivator p300 ...................................................................... 182

Inhibition of activation by PPKID4” ................................................................................ 185

Co-immunoprecipitation and pull-down of CBP ............................................................. 188
Discussion ................................................................. 191

PPKID as transcriptional activators ........................................ 191

Implications for unphosphorylated PPKID .................................. 193

Comparison to other KIX ligands ........................................... 194

Conclusions ......................................................................... 194

Experimental ......................................................................... 195

General ................................................................. 195

Materials .......................................................................... 195

Media and Buffers .......................................................... 196

Cloning of Gal4 DBD-peptide constructs ................................. 197

pPPKID4 construct .......................................................... 198

Cell culture – HEK293 cells .................................................. 200

Transcription reporter assay .................................................. 200

Transfection ...................................................................... 200

Harvesting cells .................................................................. 201

Luciferase Assay .................................................................. 201

Western Blotting .................................................................. 201

Co-immunoprecipitation ....................................................... 202

Pull-down Assay .................................................................. 203

References .......................................................................... 204
## List of Figures

### Chapter 1

**Figure 1.1** Macrocyclic scaffolds for protein surface recognition ........................................... 11

**Figure 1.2** Structure of avian pancreatic polypeptide (aPP) .................................................. 18

**Figure 1.3** Interaction between transcription coactivator CBP and transcription factor CREB ........................................................................................................................................................................... 24

**Figure 1.4** Side view of interaction between transcription coactivator CBP and transcription factor CREB ........................................................................................................................................................................... 26

**Figure 1.5** Design of PPKID Library 1 ........................................................................................ 30

**Figure 1.6** Retention of unphosphorylated library in GST-KIX selection .............................. 35

**Figure 1.7** Retention of miniature protein libraries in GST-KIX/HisKIX selection ............ 43

**Figure 1.8** Binding curves for phosphorylated control peptides ........................................... 49

**Figure 1.9** Binding curves for phosphorylated PPKID peptides ........................................... 51

**Figure 1.10** Binding curves for unphosphorylated control peptides ................................... 54

**Figure 1.11** Binding curves for unphosphorylated PPKID peptides ..................................... 56

**Figure 1.12** Binding curves for PPKID6 peptides ..................................................................... 59

**Figure 1.13** Binding curves for selected PPKID peptides with GST-KIX .......................... 62

**Figure 1.14** Competition between KID-AB$^p$ and PPKID4$^p$ or PPKID6$^U$ for binding to HisKIX ........................................................................................................................................................................................................ 65

**Figure 1.15** Competition between HisKIX-bound PPKID4$^p$flu and unlabeled PPKID4$^p$ ........................................................................................................................................................................................................ 68

**Figure 1.16** CD spectra of PPKID6$^U$ ..................................................................................... 70

**Figure 1.17** CD spectra of PPKID7$^U$ ..................................................................................... 72

**Figure 1.18** CD spectra of PPKID8$^U$ ..................................................................................... 74

**Figure 1.19** CD spectra of PPKID4$^p$ ..................................................................................... 76
Figure 1.20  CD spectra of PPKID5p ................................................................. 78
Figure 1.21  CD spectra of PPKID6p ................................................................. 80
Figure 1.22  Overview of CBP KIX variants ....................................................... 83
Figure 1.23  Close-up view of phosphoserine contacts in the KID-ABp•CBP KIX Complex ................................................................. 87
Figure 1.24  Binding curves for KID-ABp with alanine variants of CBP KIX phosphate contacts ................................................................. 89
Figure 1.25  Binding curves for PPKID4p with alanine variants of CBP KIX phosphate contacts ................................................................. 91
Figure 1.26  Close-up view of hydrophobic contacts in the KID-ABp•CBP KIX complex ................................................................. 94
Figure 1.27  Binding curves for KID-ABp with alanine variants of CBP KIX hydrophobic contacts within the binding cleft ....................................................... 96
Figure 1.28  Binding curves for PPKID4p with alanine variants of CBP KIX hydrophobic contacts within the binding cleft ....................................................... 98
Figure 1.29  Binding curves for KID-ABp with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft ....................................................... 102
Figure 1.30  Binding curves for PPKID4p with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft ....................................................... 104
Figure 1.31  Binding curves for PPKID6U with alanine variants of CBP KIX phosphate contacts ................................................................. 106
Figure 1.32  Binding curves for PPKID6U with alanine variants of CBP KIX hydrophobic contacts within the binding cleft ....................................................... 109
Figure 1.33  Binding curves for PPKID6U with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft ....................................................... 112
Figure 1.34  Side chain packing in the aPP hydrophobic core .................................. 115
Figure 1.35  Binding curves for PPKID4p α-helix variants with CBP KIX ............... 117
Figure 1.36  Binding curves for PPKID4p proline variants with CBP KIX ............ 119
Figure 2.37 Relative stabilities of the CBP KIX complexes of KID-AB\(^p\), PPKID4\(^p\) and PPKID6\(^q\) mapped on the structure of CBP KIX ........................................... 133

Chapter 2

Figure 2.1 Natural DNA-binding domains ................................................................. 157

Figure 2.2 Non-natural DNA-binding domain............................................................. 160

Figure 2.3 Plasmid map of the pAL, expression vector ............................................. 169

Figure 2.4 Schematic of transcription reporter assay ............................................... 171

Figure 2.5 Expression of Gal4 DNA-binding domain constructs .............................. 174

Figure 2.6 Transcriptional activation by PPKID peptides and KID-AB in HEK293 cells ................................................................. 179

Figure 2.7 Effect of E1A protein on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells ....................................................... 183

Figure 2.8 Effect of p300 protein on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells ....................................................... 186

Figure 2.9 Effect of PPKID4\(^p\) on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells ....................................................... 189
List of Tables

Chapter 1

Table 1.1 Binding and washing conditions for unphosphorylated GST-KIX selection .................................................................................................................. 33

Table 1.2 Occurrence of single cysteine residues in unphosphorylated GST-KIX selection .................................................................................................................. 37

Table 1.3 Amino acid sequence of selected peptides ................................................................................................................................. 39

Table 1.4 Occurrence of selected sequences for GST-KIX selection ............................................................................................................................... 40

Table 1.5 Binding and washing conditions for phosphorylated and unphosphorylated GST-KIX/HisKIX selections ................................................................................................................................. 42

Table 1.6 Selected peptides for GST-KIX/HisKIX selections ............................................................................................................................... 46

Table 1.7 HisKIX-binding affinity of selected PPKID peptides ............................................................................................................................... 47

Table 1.8 Binding affinity of PPKID and CBP KIX variants ............................................................................................................................... 85

Table 1.9 CBP KIX-binding affinity of PPKID\(^p\) variants ............................................................................................................................... 121
Chapter 1  Selection of Miniature Protein Ligands for CBP KIX
Introduction

Protein-protein interactions

Protein-protein interactions are involved in every aspect of biological processes such as cell growth, differentiation, gene expression and signal transduction. Significant efforts are focused on identifying and understanding the details of natural protein-protein interactions, which will be invaluable in elucidating the biological function of novel proteins identified through the human genome project. Additionally, this information is essential in the development of protein-protein antagonists that can be used to study biological pathways or for therapeutic benefits.

Protein-protein interfaces

The large majority of current marketable drugs target the relatively small active sites of enzymes. Targeting cell signaling and other functions controlled at protein interfaces is much more challenging because they depend on very broad protein surface interactions. To date, there is very little known about such interactions and therefore designing molecules to inhibit such surface binding events is difficult. A survey of 75 protein-protein complexes of known three-dimensional structure revealed that the typical protein-protein interface buries on average 1600 Å² of protein surface area, evenly divided between the two proteins.¹ While protein surfaces involved in protein-protein interactions are often similar in topology to the non-interacting protein surfaces, studies have found that aromatic and arginine residues occur more frequently at interfaces compared to the solvent-accessible protein surfaces. Also, the charged residues aspartic acid, glutamic acid and especially lysine are less commonly found at protein-protein
interfaces. Despite these trends, the amino acid composition of protein interfaces can vary widely,\(^1\) enabling specificity of protein-protein binding interactions. While it is clear that specific binding interfaces do exist, the driving forces for these interactions are not readily apparent. One model recognizes that the prevalence of aromatic residues at protein interfaces could dictate binding events through burial of hydrophobic regions;\(^2\) however, hydrogen bonds and salt bridges are common at protein interfaces, with an average of one hydrogen bond or salt bridge per 170 Å\(^2\) of protein interface area\(^1\) and these interactions cannot be discounted.

Alanine mutagenesis of protein interfaces is a common method for identifying contact residues important for binding affinity. Contact residues critical for binding affinity are often clustered together on the protein surface and these regions are referred to as hot spots.\(^2\) For example, structural analysis of the protein interface between human growth hormone (hGH) and its receptor (hGHbp) revealed approximately 30 side chains on each protein make contact, burying a total surface area of 2600 Å\(^2\). However, when each of these side chains on hGHbp are individually substituted with alanine, fewer than half of the mutations lead to a significant loss in binding affinity.\(^2\) In fact, a central hydrophobic region, dominated by only two tryptophan residues accounts for more than three-quarters of the binding free energy. Charged groups capable of forming salt bridges and hydrogen bonds surround this central hydrophobic region. In contrast to these key residues, several residues surrounding these charged groups have little effect on binding affinity, even though they are part of the protein-protein interface. A similar analysis of hGH revealed eight out of 31 side chains accounted for the majority of the binding energy, while over half of the residues make no significant contribution to binding
affinity.\textsuperscript{3} Even more striking, was the discovery that the region containing the most critical binding residues on hGH interacts with the corresponding hot spot region on hGHbp. This data suggests that the binding free energy between hGH and hGHbp results from a few strong interactions, rather than from the sum of all interactions at the large interface.\textsuperscript{2} Hot spot analysis has proven useful for peptide hormones and hormone receptors as well as antibodies and their respective antigen,\textsuperscript{4-10} and more recently this strategy was applied to a protein-DNA interaction\textsuperscript{11} indicating its general use.

While traditional alanine scanning is a powerful tool, it requires a large amount of time and is not well suited to study the immense number of novel protein-protein interactions discovered through genomic and proteomic efforts. As a result, a combinatorial strategy called shotgun alanine scanning\textsuperscript{12} was developed to perform alanine scanning on a large sample set. In this method, binding selections are performed using phage display libraries in which multiple sites are preferentially mutated to display either the wild type or an alanine codon. Selected library members are sequenced to determine the wild type/alanine ratio at each varied position. From these ratios, the effect of each alanine substitution is calculated in terms of the free energy change relative to the wild type ($\Delta\Delta G$). Importantly, the $\Delta\Delta G$ values obtained from shotgun alanine scanning correlate with those obtained from traditional alanine scanning. Not only do these experiments provide a wealth of information concerning the wild type protein complex, but it simultaneously examines each library member and may reveal novel protein partners with enhanced or diminished binding affinities.
**Protein surface recognition**

Disrupting protein-protein interactions presents a difficult design challenge and this challenge has been approached in a number of different ways. The success of these different methods indicates that there is not just one viable solution to the recognition of protein surfaces. The various methods described below include protein minimization, which entails finding the smallest functional epitope based on the natural protein, as well as the use of scaffolds to generate molecules that bind a protein surface.

**Protein Minimization**

As described above, only a few key residues commonly dominate protein-protein interactions. It is not necessary to recapitulate the entire binding surface rather it is only necessary to reproduce the essential features of one binding partner in a smaller form. This minimization concept has been applied in conjunction with structure-based strategies, function-based strategies or a combination of the two strategies to develop a variety of protein-protein antagonists. For example, Imperiali and coworkers used a structure-based strategy to design a zinc finger-like peptide (that does not require zinc to fold) based on the natural zinc finger Zif268.\textsuperscript{13} The resulting structure is a well-folded 23-residue peptide containing a $\alpha$-helix and $\beta$-hairpin held together by a hydrophobic core. It is likely that this zinc-less finger protein may not be functional, since there have not been reports of its use in binding studies, however, there are examples where this designed protein is used as a structural scaffold.\textsuperscript{14}

The first active small peptide mimetics were discovered through a function-based screen for an antagonist of the type 1 interleukin 1 (IL-1) receptor and a mimetic of the hormone erythropoietin.\textsuperscript{15,17} The 15-mer peptide antagonist of the IL-1 receptor was
selected through two stages of phage display. The first stage of phage display identified weak binders by fusing a library of peptides to the gene VIII coat protein of phage which displays hundreds of copies on the phage particle coat, magnifying the affinity of a single peptide; peptides that bind in the high micromolar range are tightly bound to the target in this screen. Peptides selected from the first stage of phage display were analyzed to identify a consensus binding sequence, these peptides bound IL-1 weakly with IC\textsubscript{50} values around 100 µM. Interestingly, four of five residues in the consensus sequence are consistent with residues in the natural antagonist for the IL-1 receptor. A second-generation phage display library was designed to incorporate the consensus binding sequence. These library members were displayed on the gene III coat protein of phage in a monovalent manner for a more stringent screen of potential antagonists. From this selection a 15-residue sequence was identified that has low nanomolar affinity for the IL-1 receptor.

The peptide mimetic of the hormone erythropoietin (EPO) that binds the EPO receptor was identified similarly to the IL-1 receptor antagonist\textsuperscript{16} The 14-mer cyclic peptide agonist of the EPO receptor was identified from phage display peptide libraries. Unlike the antagonist for IL-1, the consensus sequence of the selected peptides is not found in the primary sequence of EPO. Despite differences in the binding sequence, selected peptides bound the EPO receptor with high nanomolar affinity; the natural ligand binds with picomolar affinity. A crystal structure of the complex between a selected peptide and the EPO receptor suggests that the peptide and EPO share a similar binding mechanism\textsuperscript{18}.
A combination of structure-based and function-based strategies can also generate high affinity ligands for protein surfaces. First the natural protein-binding domain is reduced to only residues that make direct contact with the target protein and then phage display methods are used to restore structure and affinity to the minimized domain. This process was successful in generating a 33-residue miniprotein from the 59-residue Z-domain of Protein A, which binds to immunoglobulin G with similar affinity as the natural domain.19

Peptide Scaffolds

Peptide scaffolds are another method used to recapitulate protein surfaces. The use of peptide scaffolds approaches the challenge of recapitulating a binding epitope in the opposite direction as protein minimization. Protein minimization strategies begin with the natural protein ligand and this parent protein is whittled down to a smaller ligand that may or may not be stable and structured. On the other hand, peptide scaffolds begin with an unrelated, stably folded scaffold and attempt to incorporate a functional epitope that mimics the activity of the natural protein. Therefore, the protein minimization and peptide scaffold approaches both have inherent advantages and disadvantages.

There are a number of proteins that could serve as scaffolds that are stably folded and present protein surfaces that can suffer mutation in order to mimic its protein archetype. For example, the scorpion toxin scaffold, scyllatoxin, is stabilized by three internal disulfide bonds and presents a β-hairpin motif that is structurally equivalent to a region of CD4 that binds to the HIV-1 envelope glycoprotein (gp120). This scyllatoxin mimic of CD4 was developed through rational design followed by NMR and mutational analysis and subsequent optimization. The final molecule (CD4M9) successfully inhibits
CD4 from binding to gp120 with IC\textsubscript{50}’s ranging from 0.1 – 1.0 µM.\textsuperscript{20} Despite these complex evolution and optimization processes, CD4M9 binds gp120 about 100-fold worse than natural CD4. Further optimization of CD4M9 through mutations and incorporation of non-natural residues resulted in a significantly better molecule (CD4M33), which inhibits binding of CD4 to gp120 with IC\textsubscript{50}’s ranging from 4 – 7.5 nM, comparable to natural CD4.\textsuperscript{21}

GCN4 is another protein that can be used as a stable scaffold. The coiled coil leucine zipper domain of GCN4 was used as a scaffold for helices A and C of interleukin-4 (IL-4).\textsuperscript{22} This protein mimic was constructed using structure-based design, where the two structures are aligned to determine the best position on GCN4 to transfer IL-4 receptor binding residues. The best rationally designed peptide bound weakly to IL-4R\alpha (K\textsubscript{d} = 26 µM) compared to 1.4 nM for the native ligand IL-4. Further stabilization of the coiled coil structure with a disulfide only slightly improved affinity for IL-4R\alpha (K\textsubscript{d} = 5 µM).

These two examples illustrate an important point – it is difficult to predict how a particular mutation will affect binding. Residues that seemingly are not involved in binding can play a role in the structural presentation of the functional epitope, whereas residues that make contacts at the protein interface may have little or no contribution to binding. Therefore, rationally designed molecules often have weaker affinity for the target protein. However, the combination of rational design with a functional selection often leads to high affinity binders.
Peptoid scaffolds

In addition to native peptide scaffolds, non-natural scaffolds composed of peptoids, have been used to successfully target protein surfaces. Peptoids are N-substituted glycines that can adopt helical structure similar to polyproline type I helices. Since the substitution pattern on the amide nitrogen can be varied, this scaffold is highly adaptable to satisfy binding requirements of a variety of protein surfaces. Lim and colleagues have shown that peptoids can replace prolines in the SH3-binding motif PxxP to yield specific molecules with higher affinities (Kd = 40 nM) than natural ligands. Furthermore, the peptoid side chains can be varied to improve the selectivity among SH3 domains; the substitution of only one proline residue with an N-substituted residue resulted in a ligand with nanomolar affinity and high specificity for the Crk SH3 domain.

Lim’s approach of targeting SH3 domains with peptoids is specialized for proline-rich motifs; however, peptoid libraries have been used to select ligands for a number of other protein targets. Kodadek and colleagues synthesized a peptoid library in search of ligands that could antagonize the p53•hDM2 interaction. An initial screen yielded a peptoid that bound hDM2 with modest affinity (Kd = 37 µM). In another approach, a ligand known to bind the p53-binding domain of hDM2, chalcone 1 (Kd = 220 µM), was appended to the end of each molecule in a combinatorial library of peptoids to identify a peptoid sequence that binds to hDM2 in a non-competitive manner with chalcone 1. Indeed, the selected ligand displayed enhanced binding affinity, with an equilibrium dissociation constant of 1.3 µM. This is one example that highlights how peptoid libraries can be used to contribute binding affinity for protein surfaces.
**β-peptide scaffolds**

β-peptides are an emerging scaffold that have recently been used to target protein-protein interactions. β-peptides are homologous to α-peptides, containing an additional backbone carbon atom. Under limited conditions, they can stably fold into helices, sheets and turns, which are different from those formed by α-peptides. In one study, a β-peptide scaffold was designed that folds into a 14-helix while presenting three residues on one face of the helix that comprise the p53 functional epitope. This 10-residue β3-peptide (β53-1) is stably folded under physiological conditions and inhibits the interaction between the activation domain of p53 and hDM2 with a dissociation constant of 368 nM. This scaffold is still relatively new and its limits have not been fully explored, but it appears to be an attractive strategy for generating a variety of ligands because of its relatively small size and ability to tolerate substitutions while maintaining a stable fold.

**Macrocyclic scaffold**

Hamilton and coworkers have described various scaffolds that target the cationic surface of cytochrome c in order to disrupt its interaction with cytochrome c peroxidase (Figure 1.1). The calixarene scaffold has a bowl-shaped structure with four positions on the upper rim to attach functionality. Calixarene is thought to mimic antibody-like recognition motifs due to the large surface area it displays. When calixarene is functionalized with four identical peptide loops, a molecule that binds to cytochrome c is generated. More recent research has focused on improving synthetic methods to functionalize the calixarene scaffold with different peptide loops to improve the utility of this general scaffold for recognizing a variety of protein surfaces. Other
Figure 1.1 Macro cyclic scaffolds for protein surface recognition. (A) The calixerane scaffold for display of short peptides.\textsuperscript{32} (B) The tetraphenylporphyrin\textsuperscript{31} and (C) tetrabiphenylprophyrin scaffold designed for recognition of cytochrome c.\textsuperscript{29}
scaffolds used to bind cytochrome c include a tetraphenylporphyrin scaffold, which was used to generate a molecule that binds to cytochrome c with a $K_d$ of 20 nM and a tetrabiphenylporphyrin scaffold, which was used to identify a molecule that binds to cytochrome c with the strongest affinity of all three ($K_d = 0.67$ nM).

**Therapeutic antibodies**

The immune system is remarkably effective at protein surface recognition and as a result, antibodies bind protein surfaces with high affinity and specificity. Despite their relatively large size, antibodies are widely developed by the drug industry for therapeutic benefits. The selection of therapeutic antibodies has benefited immensely from phage-display technology, which is the most used technology to generate new human antibodies. Currently many antibodies are approved and hundreds more are in clinical development to treat cancer, autoimmune diseases and infection.

**Small molecule inhibitors**

Small molecule inhibitors for protein interfaces have lagged behind protein and peptide ligands because they lack the surface area necessary to target the large surface areas of protein interfaces. However, great strides in small molecule discovery aided by interest from pharmaceutical companies have helped to overcome many of the initial obstacles. Various methods are currently employed to identify small molecules that antagonize protein-protein interactions including high-throughput-screening (HTS) as well as non-HTS methods such as structure-based design, in silico screening and fragment-based discovery.
One of the earliest small molecule inhibitor of a protein-protein interaction was Ro26-4550, which binds the cytokine interleukin-2 (IL-2). Initially designed as a peptidomimetic of IL-2 and expected to bind to the IL-2 receptor, remarkably, this compound was found to bind IL-2 in the low micromolar range. Structural characterization of Ro26-4550 revealed that the molecule bound at a protein hot spot and inhibited the natural protein-protein interaction. This example indicates that despite the size of small molecules, they are capable of targeting the protein hot spot and disrupting the natural interaction. Ro26-4550 was used as a lead compound to design other molecules that bind to IL-2 like SP-4206, which binds to IL-2 with an IC$_{50}$ of 60 nM. SP-4206 was developed through fragment-based discovery, a process where the molecule is divided into individual components that are tested for binding to the target protein. Fragments that show any binding are then used as the starting fragment upon which to build a better inhibitor. In this case, fragment-based discovery resulted in a molecule that bound 50-fold better to the target, IL-2.

Computational methods are important in the discovery of small molecule inhibitors and aid non-HTS methods including structure-based design, in silico screening and fragment-based discovery. These computational methods are used at all stages of drug discovery, particularly in de novo design, where inhibitors of protein-protein interactions are designed from purely creative efforts, and molecular docking using in silico library screening, where libraries of potential ligands are positioned in the binding site and scored for potential activity. Both of these methods require a structure of the target binding site, which is a limiting factor. However, once the binding surface is
identified with biochemical data, programs exist that can design and/or screen ligands for binding site compatibility.

As the number of small molecule inhibitors has increased, so have their applications. There are a number of examples from the Schreiber laboratory in which protein-binding small molecules were used to understand a biological pathway. For example, the discovery of FK506-binding protein (FKBP12) led to the elucidation of the calcium-calcineurin-NFAT signaling pathway.\(^{39}\) Additionally, the discovery of the small molecule rapamycin, which simultaneously binds FKBP12 and FRAP led to a better understanding of the nutrient signaling network, leading to insight on the origins of type 2 diabetes.\(^{40}\) These examples demonstrate the wide utility of small molecules for targeting protein surfaces, even those that are not in the active site of enzymes.

**Alpha helix stabilization**

In phage display selections, peptides identified to bind proteins are often structured; this suggests that pre-organization of peptides may be a requirement for high-affinity binding to protein surfaces.\(^{41}\) However, small (<50 residues), monomeric, structured protein domains that lack metal binding sites and disulfides are rare in nature.\(^{42}\) As a result, significant efforts have focused on promoting pre-organization in peptide scaffolds through stabilization of secondary structure elements. There are many methods to stabilize an \(\alpha\)-helix including disulfide bonds, lactam bridges and cross-linking.

**Disulfide and other covalent bonds**

There are several examples where introducing disulfide bonds into protein ligands has increased their stability, and in most cases, resulted in higher affinity for its target
protein.\textsuperscript{22,42} For example, Z38 is a 38-residue, minimized version of the Z-domain of Protein A discussed previously. Z38 exhibits significant $\alpha$-helical character, but it is not thermally stable ($T_m < 10\, ^\circ C$). However, removing five residues at the N-terminus and adding a disulfide bond resulted in a molecule (Z34C) with increased stability; the $T_m$ increased by more than a 40 °C.\textsuperscript{42} An important effect is that the affinity of Z34C for immunoglobulin increased 9-fold over Z38 to 20 nM through this structural stabilization.

In a similar strategy to disulfide bridges, it is possible to incorporate new covalent linkages through incorporation of lactam bridges and hydrocarbon cross-links. Lactam bridges are created through amide bond formation between the terminal amine of lysine and the carboxyl groups of aspartate and glutamate at $i$ and $i+4$ positions,\textsuperscript{43} or between two glutamate residues coupled with a diaminoalkane at positions $i$ and $i+7$.\textsuperscript{44} Molecules stabilized by lactam bridges often have similar structures and affinities as though the molecule was stabilized with a disulfide bond. Similarly, hydrocarbon cross-linking through the incorporation of two unnatural amino acids at $i$ and either $i+4$ or $i+7$ positions was shown to stabilize $\alpha$-helicity and increase metabolic stability in peptides.\textsuperscript{45} The unnatural amino acids, which are $\alpha$-methylated with olefinic side chains, can be connected via ruthenium-catalyzed olefin metathesis to generate a cross-link spanning one or two turns of the $\alpha$-helix. These approaches, however, have limited utility because of the chemical modifications necessary to promote the bond formation.

\textit{Non-peptidic scaffolds}

Hamilton and coworkers reported a new family of $\alpha$-helix mimetics based on a functionalized terphenyl scaffold.\textsuperscript{46} They found that the $3,2',2''$-substituents on a terphenyl in a staggered conformation have similar angular relationships and distances as
the i, i+3 and i+7 groups of an \( \alpha \)-helix. A terphenyl scaffold was used to design a mimic of the \( \alpha \)-helical domain of smooth muscle myosin light-chain kinase (smMLCK), which binds calmodulin. Terphenyl analogues were generated that bind calmodulin in the nanomolar range (IC\(_{50} = 9-800 \) nM) comparable to a peptide corresponding to the \( \alpha \)-helical domain of smMLCK (IC\(_{50} = 80 \) nM). Other \( \alpha \)-helix mimetics designed in the Hamilton lab include a polyamide foldamer made up of a functionalized trispyridylamide scaffold that binds Bcl-X\(_L\) with low micromolar affinity.\(^{47}\)

**Avian Pancreatic Polypeptide (aPP)**

All of the above stabilization methods use non-natural scaffolds or require post-translational modification; a completely encodable scaffold with inherent \( \alpha \)-helicity would have significant advantages in terms of enhanced binding affinities through pre-organization. Avian pancreatic polypeptide (aPP) is a member of the PP fold peptide family of hormones. Pancreatic polypeptide is found in the endocrine pancreas and its release is stimulated by food intake, indicating it may act as a satiety factor. Pancreatic polypeptide also has effects on gut motility and gastrin-like actions.\(^{48}\) Mammalian and avian pancreatic polypeptides have close sequence homology despite differing at 20 of 36 positions. The aPP structure consists of a polyproline type II helix connected to an \( \alpha \)-helix by a short turn (Figure 1.2). The stability of aPP (\( T_m = 60^\circ C \)) is derived from a hydrophobic core that is created when the polyproline helix folds back onto the \( \alpha \)-helix.\(^{49}\) In theory, the stable aPP scaffold can be used as a mimic for any \( \alpha \)-helix protein to antagonize protein-DNA or protein-protein interactions.
Figure 1.2 Structure of avian pancreatic polypeptide (aPP).\textsuperscript{49} The polyproline type II helix and $\alpha$-helix pack against each other forming a hydrophobic core that stabilizes the molecule.
C-terminus

Hydrophobic Core

N-terminus
Protein Grafting

The Schepartz laboratory has described a general solution for macromolecular recognition, called protein grafting.\textsuperscript{50} In protein grafting, the residues that comprise a functional epitope are grafted onto the solvent-exposed $\alpha$-helical face of the small yet well-folded protein aPP. It is hypothesized that pre-organization of the binding epitope will lead to high affinity ligands. This procedure, often in combination with molecular evolution, identifies miniature protein ligands with high affinity and specificity for macromolecular targets.

Using the $\alpha$-helix of aPP to target DNA

The protein grafting strategy was first used to design a protein displaying the DNA-contact residues of GCN4.\textsuperscript{50} GCN4, a member of the class of bZIP transcription factors, binds the DNA sequence half-site CRE (hsCRE) with an $\alpha$-helix. The rationally designed molecule, PPBR4, bound the hsCRE with high affinity ($K_d = 1.5$ nM) at 4 °C and exhibited greater sequence specificity than the bZIP element of GCN4. However, PPBR4 does not retain the structure of the aPP scaffold. To improve the structure of PPBR4, while retaining or improving its ability to bind hsCRE, a phage display library based on PPBR4 was generated\textsuperscript{51} to identify library members that retain both structure and function. Residues were varied on the polyproline helix to optimize packing against the chimeric alpha helix. After 3 rounds of phage display, selected sequences bound hsCRE as well as, or better than PPBR4. One sequence, p007, bound the hsCRE 100 times better than PPBR4 and retains this affinity at 25 °C ($K_d = 1.6$ nM), a temperature where PPBR4 and the bZIP element of GCN4 do not bind the hsCRE. p007 also
discriminated between several duplex oligonucleotides where two of the five base pairs were mutated. The structure of p007, as determined by NMR, is very similar to the structure of aPP, indicating the stabilization of the α-helical structure is imperative for high affinity.

In a second example, the protein grafting strategy was used in conjunction with phage display to identify ligands that mimic the DNA-binding domain of the transcription factor Q50K engrailed homeodomain. The strategy resulted in the peptide PPeng4, which bound the DNA containing the QRE target site with high affinity ($K_d = 17$ nM) and specificity.$^{52}$ More importantly, this was the first example where it was demonstrated that pre-organization of the functional epitope onto a well-folded α-helix can compensate for an incomplete set of DNA-binding residues.

*Using the α-helix of aPP to target proteins*

With the success of the aPP scaffold to generate ligands for DNA, this approach was applied to protein targets. Once again, the protein grafting strategy combined with functional selection using phage display was used to design and evolve a miniature protein that binds the protein surface of Bcl-2 and Bcl-X$_L$, two proteins involved in apoptosis. The miniature protein, PPBH3-1, bound Bcl-2 with 100 times higher affinity than the natural ligand Bak and displays a preference for Bcl-X$_L$ similar to Bak.$^{53}$ Further selection with a phage display library showed the same strategy can be used to select for a ligand with the reverse specificity; ligands were identified that have a preference for Bcl-2 over Bcl-X$_L$. $^{54}$
Using the PPII helix of aPP to target proteins

Most recently, the polyproline helix portion of the aPP scaffold was used to target proteins. Many proteins involved in cellular signaling recognize proline-rich sequences including SH3, WW and EVH1 domains. The target protein, Mena, is a mammalian protein containing an EVH1 domain and involved in regulation of actin filament dynamics. EVH1 domains interact with target proteins containing proline-rich motifs including the ActA protein of pathogen *Listeria monocytogenes*. Through rational design, the functional epitope of ActA was incorporated into the polyproline region of aPP. The resulting miniature protein, pGolemi, specifically binds Mena over other EVH1 domains Vasp and Evl. In addition, pGolemi was shown to disrupt actin motility of *L. monocytogenes* in mammalian cells.

Extension of protein grafting to flat protein surfaces

The protein grafting strategy was successful in generating high affinity ligands for Bcl-2 and Bcl-XL, however, these proteins contain a deep cleft where the α-helix inserts into the surface of the protein. As discussed above, many protein-protein interactions occur between flat surfaces on the proteins. Here we use the interaction between transcription factor CREB and transcriptional coactivator CBP to show that protein grafting can also be used to target flat surfaces. Additionally, the CREB•CBP interaction provides another challenge in the necessity to incorporate a post-translational modification step into the selection. CREB requires phosphorylation on a serine residue in order to bind CBP, therefore a phosphorylation step will be incorporated into the phage display selection.
The transcription factor CREB

The transcription factor CREB (cyclic AMP response element-binding protein) contains a DNA-binding domain, which binds to a conserved cAMP responsive element (CRE), and two activation domains - a glutamine-rich domain (Q2) and a kinase-inducible domain (KID). CREB binds to the palindromic CRE sequence as a dimer, similar to other bZIP proteins. The Q2 region acts as a constitutive activation domain, binding to TAF130 of the TFIID complex, whereas the KID requires phosphorylation at Ser133 to activate transcription. CREB is required for many cellular processes including proliferation, differentiation and adaptive responses. CREB KID functions in conjunction with a transcriptional coactivator CBP (CREB-binding protein).

KID-KIX interaction

The kinase-inducible domain of CREB is a downstream target of the cyclic AMP (cAMP) signaling pathway. Levels of cAMP are increased in response to extracellular signals. Increased levels of cAMP activate protein kinase A (PKA) by binding to its regulatory subunit. This frees the catalytic subunit of PKA, which can then phosphorylate CREB at Ser133. Phosphorylation at this residue increases the affinity of CREB for the transcription coactivator CBP, resulting in increased expression of CREB-dependent genes through CBP's interaction with the RNA polymerase complex.

The CBP-binding activity of CREB has been localized to the KID, which binds to a shallow hydrophobic groove formed by α-helices 1 and 3 of the KIX domain within CBP (Figure 1.3). The KID domain (residues 121-146) is made up of two α-helices, αA (residues 121-129), and amphipathic αB (residues 133-146). As shown in Figure 1.4, CREB KID binds to essentially a flat surface on KIX and the interaction
**Figure 1.3** Interaction between transcription coactivator CBP and transcription factor CREB. The CBP KIX domain (gray surface) forms a shallow hydrophobic groove in which the CREB KID (ribbon) binds. Binding residues on CREB are shown explicitly. The CREB KID is made up of two helices A (orange) and B (red). The majority of the contacts between CREB and CBP are made with helix B.
Figure 1.4 Side view of interaction between transcription coactivator CBP and transcription factor CREB. The CBP KIX domain (gray surface) forms a shallow hydrophobic groove in which the CREB KID (ribbon) binds. Binding residues on CREB are shown explicitly. The A and B helices of CREB KID are in orange and red, respectively.
completely buries only one residue. The KID is unstructured when not bound to CBP, however it undergoes a random coil to helix transition upon binding CBP to form two α-helices; each helix contacts a different region of the CBP KIX surface.\textsuperscript{60} Most of the KID residues important for binding are contained within the B helix. Four KID residues critical for KIX binding are Y134, I137, L138 and L141. Three residues, L128 (A helix), A145 and P146 also make hydrophobic contacts, but contribute less to the free energy of complex formation. In addition, phosphorylated S133 forms a hydrogen bond with Y658 and a salt bridge with K662 of KIX. The interaction between KID and KIX buries approximately 1200 Å\textsuperscript{2} of solvent-accessible surface, over 950 Å\textsuperscript{2} of this surface area comes from contacts with the B helix of KID.\textsuperscript{60}

**Synthetic KIX-binding activation domains**

The first synthetic KIX ligands were selected from a completely degenerate 8mer peptide phage display library.\textsuperscript{61} These ligands bound CBP KIX with equilibrium dissociation constants ranging from 16 to 55 µM and activated transcription 10 to 40-fold. Analysis of the binding footprint of KBP 2.20, the ligand with highest affinity, on CBP KIX revealed that the peptide bound an overlapping, yet minimal surface of KIX as compared to phosphorylated CREB or unphosphorylated KIX ligand c-Myb. Moreover, KBP 2.20 appeared to penetrate further into the hydrophobic groove. However, one disadvantage of these ligands is that they are unstructured and hence more susceptible to proteolysis. Our approach described below takes into account the natural ligand and provides a structured scaffold leading to high affinity KIX ligands that have the propensity to fold, yet are still relatively small in size.
Design of PPKID Library 1

The design of a CBP KIX-binding miniature protein (PPKID) library was based on the alignment of the α-helix of aPP and helix B of the CREB KID domain (Figure 1.5). Mutagenesis studies have determined that most (though not all) of the residues that comprise the CBP KIX-binding epitope of CREB KID\textsuperscript{p} are located in helix B,\textsuperscript{50,62} therefore only residues from helix B were included in the miniature protein library. Four hydrophobic residues from CREB KID (Tyr134, Ile137, Leu138, Leu141) contribute significantly to the free energy of KID\textsuperscript{p}•KIX complex formation. The PPKID library contains three of these four residues (Ile137, Leu138, Leu141), and a conservative mutation of the fourth from Tyr to Phe, which in the context of CREB KID\textsuperscript{p} has no effect on CBP KIX binding.\textsuperscript{63} This mutation was included, along with the complete recognition site for protein kinase A (PKA; Arg130, Arg131, Ser133), to promote phosphorylation of the miniature protein library in vitro, if so desired. In the context of CREB KID\textsuperscript{p}, the Tyr to Phe mutation lowers the $K_m$ for PKA-mediated phosphorylation by five-fold.\textsuperscript{63} The structural scaffold of the α-helical portion of the library was provided by five of seven residues (Leu17, Phe20, Leu24, Tyr27, Leu28) from the aPP α-helix that contribute to the hydrophobic core.\textsuperscript{49} Based on success in our laboratory using a similar approach to improve DNA-binding miniature proteins,\textsuperscript{51} the five residues from the polyproline helix of aPP known to participate in hydrophobic core formation (Pro2, Gln4, Pro5, Tyr7, Pro8) were varied to all 20 amino acids in the PPKID library. The expectation was that the CBP KIX-binding epitope on the α-helix would guide all library members to the CBP KIX surface, and the functional selection would identify those library members with increased CBP KIX affinity derived from packing of the polyproline helix against the
**Figure 1.5** Design of PPKID Library 1. (A) Schematic of protein grafting. In the KID-AB\textsuperscript{p}•CBP KIX complex, the residues along helix B that comprise the functional epitope are in red, and the PKA recognition site used to direct the PKA-dependent phosphorylation of serine 18 is in green. These residues are displayed on the solvent-exposed face of the aPP α-helix while residues on the polyproline helix are randomized to form a miniature protein library. Library members that bind to CBP KIX are selected through phage display. (B) Alignment of the α-helix sequences of aPP and KID-B\textsuperscript{p} (residues 130-146). Residues along the α- and PPII-helices that contribute to formation of the aPP hydrophobic core are shown in blue, residues important for binding CBP KIX are in red, and residues that comprise the PKA recognition site are in green. Residues randomized in the library are indicated by (X).
CREB KID\(^p\) • CBP KIX complex

**A**

- CREB KID\(^p\) • CBP KIX complex
- KID\(^p\) functional epitope
- aPP scaffold
- miniature protein library

**B**

<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>aPP</td>
<td>GPSQPTYPGDDAPVEDLIRFYDNLQQLNLYVV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KID(^p) (Helix B)</td>
<td>RRPSYRKILNDLSSDAP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>GXSXTXGDDAPVRLFSYYILDLYLDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
otherwise exposed face of the bound α-helix. A $5 \times 10^7$-member library of miniature proteins based on this design was generated by Dr. Stacey Rutledge\textsuperscript{64} for use in phage display selection experiments.

Results

Experimental goals

The primary objective of this study was to use the PPKID Library 1 with phage display to select for both phosphorylated and unphosphorylated miniature proteins based on CREB KID, thus verifying that the protein grafting and molecular evolution strategy developed in our laboratory can be applied to shallow protein surface targets and post-translational modifications can be incorporated into the selection. The next goal was to characterize the selected miniature proteins using 1) fluorescence polarization to determine their affinity for CBP KIX and its variants, and 2) circular dichroism and mutagenesis to investigate their secondary structure.

Unphosphorylated phage display selection against GST-KIX

Nine rounds of panning were carried out on the unphosphorylated library according to the conditions shown in Table 1.1. Rounds 1-5 were performed by Stacey Rutledge.\textsuperscript{64} Briefly, a round consists of amplifying bacteria containing the phagemid library vector to mid-log phase, the bacteria are then infected with helper phage to produce phage-peptide particles. The phage-peptide particles are then added to a microtiter plate coated with the target protein, in this case a GST-fusion of CBP KIX (GST-KIX). After several cycles of washing the microtiter plate, phage that remain
Table 1.1 Binding and washing conditions for unphosphorylated GST-KIX selection

<table>
<thead>
<tr>
<th>Round</th>
<th>Target protein</th>
<th>Temperature (°C)</th>
<th>Number of washes</th>
<th>Wash length (min)</th>
<th>Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(^a)</td>
<td>GST-KIX</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>TBST</td>
</tr>
<tr>
<td>2(^a)</td>
<td>GST-KIX</td>
<td>4</td>
<td>10</td>
<td>5</td>
<td>TBST</td>
</tr>
<tr>
<td>3(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>2</td>
<td>TBST</td>
</tr>
<tr>
<td>4(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST</td>
</tr>
<tr>
<td>5(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST</td>
</tr>
<tr>
<td>6(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST(^b)</td>
</tr>
<tr>
<td>7(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST(^b)</td>
</tr>
<tr>
<td>8(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST + 5 mM DTT</td>
</tr>
<tr>
<td>9(^a)</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST + 5 mM DTT</td>
</tr>
</tbody>
</table>

\(^a\) Rounds 1-5 were performed by Dr. Stacey Rutledge.
\(^b\) Rounds 6B and 7B contained 5 mM DTT in the buffer.

In rounds 1-9, GST-KIX (30 nM) was immobilized on glutathione-coated microtiter wells.
bound to the plate are eluted using a low pH buffer. Bacteria are then infected with the eluted (output) phage to determine the retention rate of the library by comparing to the amount of input phage. Output phage are used to start the next round of panning. Percent retention of the library and retention of the library as compared to a negative control (phage expressing aPP) is shown in Figure 1.6. Retention decreases at round 3 due to the increase in selection temperature from 4°C to 25°C.

A representative group of clones from each round (approximately 20) were sequenced beginning at round 6. Many clones sequenced in rounds 6-9 contained single cysteine residues, suggesting that the selected peptides were binding to the target protein GST-KIX through formation of disulfide bonds. The KIX domain does not contain any cysteine residues, however the GST-tag contains four cysteine residues therefore peptides can be retained on the plate by forming a disulfide bond with the GST portion of the protein instead of binding the target domain KIX. Various studies have shown that in random peptides expressed by M13 within pIII, there is a strong negative selection against odd numbers of cysteine residues. The occurrence of cysteine residues in selected peptides are shown in Table 1.2. Addition of dithiothreitol (DTT), a reducing agent, to the binding buffer starting at round 8 did not alleviate this problem and there was not a decrease in retention when DTT was added to the binding buffer. Therefore, selections were repeated starting at round 6 including 5 mM DTT in the binding buffer (round 6B-7B). In rounds 6B and 7B, approximately 25% of the sequences contained cysteine residues compared to 40-60% in previous rounds without DTT.

One sequence – PPKID6U – was selected in rounds 6-9, 6B and 7B. PPKID6U contains an isoleucine at position 2, tryptophan at position 4, proline at position 5,
Figure 1.6 Retention of unphosphorylated library in GST-KIX selection. (A) Percent retention of unphosphorylated library. Input and output titers were used to calculate percent retention, defined as (output titer / input titer) x 100. (B) Retention of unphosphorylated library as compared to retention of aPP. Fold retention was calculated by dividing the percent retention of the library by the percent retention of phage displaying aPP.
Table 1.2 Occurrence of single cysteine residues in unphosphorylated GST-KIX selection

<table>
<thead>
<tr>
<th>Round</th>
<th>Cysteine-containing sequences</th>
<th>Total sequences</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>11</td>
<td>20</td>
<td>0.55</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>18</td>
<td>0.39</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>20</td>
<td>0.60</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>8</td>
<td>0.50</td>
</tr>
<tr>
<td>6B</td>
<td>5</td>
<td>19</td>
<td>0.26</td>
</tr>
<tr>
<td>7B</td>
<td>4</td>
<td>19</td>
<td>0.21</td>
</tr>
<tr>
<td>7C</td>
<td>5</td>
<td>16</td>
<td>0.31</td>
</tr>
<tr>
<td>8C</td>
<td>4</td>
<td>14</td>
<td>0.29</td>
</tr>
<tr>
<td>9C</td>
<td>8</td>
<td>18</td>
<td>0.44</td>
</tr>
</tbody>
</table>
phenylalanine at position 7 and a glutamic acid at position 8 (Table 1.3). In addition a second sequence, PPKID8U, with randomized residues LWKDP at postions 2,4,5,7 and 8, was also selected frequently (Table 1.4). While no consensus was reached in this selection, certain residues were preferred at particular positions. Hydrophobic residues were preferred at position 2; alanine, isoleucine or leucine represented the majority of residues selected at this position. Tryptophan was selected at varying positions in nearly every sequence that did not contain a cysteine. The most popular position for tryptophan was position 4, however when a tryptophan was not selected in position 4, it was almost always selected in position 8. In position 5, proline and charged residues were selected, mostly lysine, and to a lesser extent arginine and aspartic acid. In position 7 or 8 there was a slight preference for aromatic or negatively charged residues.

Selection 1 panning with HisKIX

In order to remove bias for GST-binding peptides and the uncharacteristically high selection of single cysteine residues, the PPKID library was panned against the target protein HisKIX. Starting with the output of round 6 from the GST-KIX selection, three additional rounds of panning were completed (rounds 7C-9C). The output of round 6 was chosen as a starting point because it was the first round where PPKID6 was selected. Overall, retentions were much lower than those seen with GST-KIX. This would be expected if the library members were biased toward peptides that bound the GST-fusion part of the protein. By round 9C, the library was retained 1.6 fold over aPP. In addition to the negative control, aPP, a positive control, Myb, was panned in parallel for these three rounds. Myb is a transcription factor that binds KIX with a Kd of 2 µM. Comparing the library retention to the retention of this positive control revealed improved
Table 1.3 Amino acid sequence of selected peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID4</td>
<td>GSPQPTYPGDDAPVRRLSFFYILLDLYLDAP</td>
<td>U and P</td>
</tr>
<tr>
<td>PPKID5</td>
<td>GLSWPTYHGDADVRRLSFFYILLDLYLDAP</td>
<td>U and P</td>
</tr>
<tr>
<td>PPKID6</td>
<td>GISWPTFEGDADVRRLSFFYILLDLYLDAP</td>
<td>U only</td>
</tr>
<tr>
<td>PPKID7</td>
<td>GLSWKTDGGDDAPVRRLSFFYILLDLYLDAP</td>
<td>U only</td>
</tr>
<tr>
<td>PPKID8</td>
<td>GLSPYTEWGDADPVRRLSFFYILLDLYLDAP</td>
<td>U only</td>
</tr>
</tbody>
</table>

Selected residues are shown in red.
U - unphosphorylated selection, P - phosphorylated selection
In phosphorylated selection, S18 is phosphorylated.
<table>
<thead>
<tr>
<th></th>
<th>Round 6</th>
<th>Round 7</th>
<th>Round 8</th>
<th>Round 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID6</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PPKID7</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9</strong></td>
<td><strong>11</strong></td>
<td><strong>8</strong></td>
<td><strong>4</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Round 6B</th>
<th>Round 7B</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>PPKID7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14</strong></td>
<td><strong>15</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Round 7C</th>
<th>Round 8C</th>
<th>Round 9C</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID6</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>PPKID7</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11</strong></td>
<td><strong>10</strong></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>
binding by the library. The fold retention of the PPKID library compared to Myb improved from 0.17 to 1.3 between rounds 7C-9C. Surprisingly, after three rounds of panning with HisKIX, peptides containing a single cysteine were still selected. It is possible that the library was biased in early rounds due to selection for the GST-tag and more rounds of panning will be needed to eliminate this bias. PPKID6 was selected in all rounds with the HIS-KIX target protein whereas PPKID8 was selected in rounds 7C and 8C, but not in round 9C.

**Phosphorylated and unphosphorylated phage display selection using HisKIX/GST-KIX**

In light of evidence that selected peptides were binding the GST-tag rather than the target protein alone, a new set of selections were performed using the PPKID Library 1. Starting with the output of round 1, HIS-KIX and GST-KIX were alternated as the target protein every other round, using the same binding and washing conditions for two rounds before increasing stringency (Table 1.5). Alternating the target proteins will minimize selection of peptides that bind preferentially to either the GST- or His-tag. In addition 5 mM DTT was included in all rounds with GST-KIX to discourage formation of disulfide bonds with the cysteine residues in GST. Two separate selections (phosphorylated and unphosphorylated) were panned in parallel. The phage in the phosphorylated selection were treated with 2500 units PKA for 2 hours to phosphorylate the serine residue at position 18 of the phage displayed peptide.

Figure 1.7 shows the absolute retention of the library at each round. There is a large decrease in retention at round 6 due to an increase in wash length from 2 to 5 minutes. In general retention is higher in rounds with GST-KIX as the target protein. While this increase in retention may be due to affinity for the GST-tag it can also be due
GST-KIX (30 nM) was immobilized on glutathione-coated microtiter wells and HisKIX (100 nM) was immobilized on Ni-NTA-coated microtiter wells.

Table 1.5 Binding and washing conditions for phosphorylated and unphosphorylated selections

<table>
<thead>
<tr>
<th>Round</th>
<th>Target protein</th>
<th>Temperature (°C)</th>
<th>Number of washes</th>
<th>Wash length (min)</th>
<th>Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>GST-KIX</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>TBST</td>
</tr>
<tr>
<td>2</td>
<td>HisKIX</td>
<td>4</td>
<td>10</td>
<td>5</td>
<td>PBST</td>
</tr>
<tr>
<td>3</td>
<td>GST-KIX</td>
<td>4</td>
<td>10</td>
<td>5</td>
<td>TBST&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4</td>
<td>HisKIX</td>
<td>25</td>
<td>10</td>
<td>2</td>
<td>PBST</td>
</tr>
<tr>
<td>5</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>2</td>
<td>TBST&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>6</td>
<td>HisKIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>PBST</td>
</tr>
<tr>
<td>7</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>8</td>
<td>HisKIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>PBST</td>
</tr>
<tr>
<td>9</td>
<td>GST-KIX</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>TBST&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*Round 1 was performed by Dr. Stacey Rutledge.*

*Buffer contains 5 mM DTT.*

<sup>a</sup> GST-KIX (30 nM) was immobilized on glutathione-coated microtiter wells and HisKIX (100 nM) was immobilized on Ni-NTA-coated microtiter wells.
Figure 1.7 Retention of miniature protein libraries in GST-KIX/HisKIX selection. (A) Percent retention of unphosphorylated library. (B) Percent retention of phosphorylated library. Input and output titers were used to calculate percent retention, defined as (output titer / input titer) x 100.
A

Unphosphorylated Library

Percent Retention

GST  His  GST  His  GST  His  GST  His  GST  His  GST  His  GST

Round #

B

Phosphorylated Library

Percent Retention

GST  His  GST  His  GST  His  GST  His  GST  His  GST

Round #

.22
to the fact that rounds with GST-KIX followed rounds with HIS-KIX of the same stringency. Between rounds 7 and 9, retention of the phosphorylated and unphosphorylated libraries increased by 7.2 and 4.3-fold.

A table of selected sequences from both the unphosphorylated and phosphorylated libraries is shown in Table 1.6. Four peptides, PPKID4, 5, 7 and 8, appeared multiple times in rounds 5-9 of the unphosphorylated selection; two of these peptides (PPKID4 and 5) appeared in both the phosphorylated and unphosphorylated selections. Alternating the target protein reduced non-specific binding and, as a result, cysteine residues were not prevalent in the selected sequences (10% of selected peptides from the unphosphorylated library contained single cysteine residues and there were no occurrences in the phosphorylated library). Interestingly, PPKID4 contains aPP-derived residues in all randomized positions. Phosphorylated ligands PPKID4 and PPKID5 contain identical residues at two of the randomized positions, 5 (Pro) and 7 (Tyr), but otherwise the selected residues are not conserved. In the unphosphorylated selection, four of five randomized positions (2, 4, 5, and 7) in peptides PPKID 4-8 approach consensus; Leu or Ile was selected at position 2, Trp at position 4, Pro at position 5, and aromatic or negatively charged residues at position 7.

**Affinity of PPKID and control peptides for HisKIX**

PPKID peptides were synthesized with a Gly-Val linker and a Cys residue on the C-terminus that was labeled with acetamidofluorescein. The affinity of each labeled peptide for HisKIX was measured by equilibrium fluorescence polarization (Table 1.7). To compare the PPKID peptides to the natural ligand, two control peptides were synthesized; KID-AB and KID-B. Peptide KID-AB comprises the full-length CREB
Table 1.6 - Selected peptides for GST-KIX/HisKIX

**Unphosphorylated Selection**

<table>
<thead>
<tr>
<th></th>
<th>Round 7</th>
<th>Round 8</th>
<th>Round 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PPKID5</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PPKID7</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>PPKID8</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>15</strong></td>
<td><strong>12</strong></td>
<td><strong>16</strong></td>
</tr>
</tbody>
</table>

**Phosphorylated Selection**

<table>
<thead>
<tr>
<th></th>
<th>Round 5</th>
<th>Round 6</th>
<th>Round 7</th>
<th>Round 8</th>
<th>Round 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID4</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PPKID5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11</strong></td>
<td><strong>13</strong></td>
<td><strong>17</strong></td>
<td><strong>12</strong></td>
<td><strong>16</strong></td>
</tr>
</tbody>
</table>
Table 1.7 HisKIX-binding affinity of selected PPKID peptides

<table>
<thead>
<tr>
<th>Phosphorylated Selection</th>
<th>$K_d^{PPKID^P}$ (nM)</th>
<th>$K_d^{PPKID^U}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID4$^P$</td>
<td>515 ± 44</td>
<td>12.1 ± 2.4</td>
</tr>
<tr>
<td>PPKID5$^P$</td>
<td>534 ± 31</td>
<td>6.6 ± 2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unphosphorylated Selection</th>
<th>$K_d^{PPKID^P}$ (nM)</th>
<th>$K_d^{PPKID^U}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPKID6$^U$</td>
<td>624 ± 49</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>PPKID6 S18E</td>
<td></td>
<td>10.9 ± 2.0</td>
</tr>
<tr>
<td>PPKID7$^U$</td>
<td>2.3 ± 0.2</td>
<td></td>
</tr>
<tr>
<td>PPKID8$^U$</td>
<td>3.1 ± 0.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control peptides</th>
<th>$P: K_d$</th>
<th>$U: K_d$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KID-AB</td>
<td>562 ± 41 nM</td>
<td>&gt;116</td>
</tr>
<tr>
<td>KID-B</td>
<td>51.6 ± 4.0 μM</td>
<td>&gt;297</td>
</tr>
<tr>
<td>Peptide C</td>
<td>2.4 ± 0.2 μM</td>
<td>21.5 ± 2.6</td>
</tr>
</tbody>
</table>
KID (residues 119-148) whereas peptide KID-B is made up of only the B helix of CREB KID (residues 130-148). KID-AB^p binds HisKIX with high affinity (K_d = 562 ± 41 nM) at 25 °C (Figure 1.8). This value is lower than previously reported equilibrium dissociation constants for similar KID^p•KIX complexes (3.1 µM to 9.7 µM). These differences may be due to the fact that the dissociation constants in each of these studies were measured by a number of techniques (though not fluorescence polarization), using slightly different buffers and KID^p and KIX constructs in each study. The dissociation constant of KID-B^p for HisKIX (K_d = 51.6 ± 4.0 µM) is comparable to the dissociation constant (K_d = 80 µM) reported for the KID(129-149)^p•KIX complex measured by isothermal titration calorimetry. The binding affinity of an additional control peptide was also measured for HisKIX. The phosphorylated (C^p) control peptide consists of the chimeric α-helical portion of the PPKID peptides (residues 15-33) without the polyproline helix or turn region. Because residues were not varied in the α-helix portion of the library, these control peptides will allow us to compare the contribution to CBP KIX-binding affinity of residues in the α-helix derived from aPP and residues in the randomized region of the PPKID library, which include the putative polyproline helix and turn regions. Any enhanced affinity compared to the control peptides can be attributed to the selected residues. The peptide C^p binds HisKIX with a K_d of 2.4 ± 0.2 µM, which represents a greater than 20-fold increase in affinity (ΔΔG = -1.8 kcal•mol^{-1}) compared to the CBP KIX-binding affinity of KID-B^p (Figure 1.8). This peptide binds surprisingly well, suggesting that the selected residues do not contribute much to KID-B^p’s affinity for CBP KIX. PPKID4^p and PPKID5^p exhibit high affinity for HisKIX (K_d = 515 ± 44 nM and 534 ± 31 nM) (Figure 1.9). Therefore, the turn and polyproline helix
Figure 1.8 Binding curves for phosphorylated control peptides. The affinity of KID-AB\(p\), KID-B\(p\) and peptide C\(p\) for HisKIX were measured by equilibrium fluorescence polarization. KID-AB\(p\) and KID-B\(p\) data were collected by Stacey Rutledge. The polarization of fluorescein-labeled peptide (peptide\(p\)Flu) was monitored as a function of HisKIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID\(p\)Flu bound values using the \(P_{\text{min}}\) and \(P_{\text{max}}\) values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID\(p\)Flu bound values to equilibrium binding equation 1.2. The \(K_d\)'s derived from this analysis for the peptide\(p\)Flu•HisKIX interaction under these conditions are 562 ± 41 nM (KID-AB\(p\)), 51.6 ± 4.0 \(\mu\)M (KID-B\(p\)) and 2.4 ± 0.2 \(\mu\)M (peptide C\(p\)).
Fraction of Peptide Bound

- KID-AB$^P$
- KID-B$^P$
- Peptide C$^P$

[Graph showing binding curves for KID-AB$^P$, KID-B$^P$, and Peptide C$^P$ against [HisKIX] (M).]
Figure 1.9 Binding curves for phosphorylated PPKID peptides. The affinity of PPKID$^{4p}$ and PPKID$^{5p}$ for HisKIX were measured by equilibrium fluorescence polarization. The KID-AB$^{p}$ binding curve is shown for comparison. The polarization of fluorescein-labeled peptide (peptide$^{p}$Flu) was monitored as a function of HisKIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID$^{p}$Flu bound values using the $P_{\text{min}}$ and $P_{\text{max}}$ values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID$^{p}$Flu bound values to equilibrium binding equation 1.2. The $K_d$’s derived from this analysis for the peptide$^{p}$Flu•HisKIX interaction under these conditions are 515 ± 44 nM (PPKID$^{4p}$), 534 ± 31 nM (PPKID$^{5p}$) and 562 ± 41 nM (KID-AB$^{p}$).
regions (including selected residues) of the PPKID4p and PPKID5p peptides contribute a modest 0.9 kcal•mol⁻¹ to the free energy of complex formation with CBP KIX.

The HisKIX-binding affinities of unphosphorylated versions (denoted by a superscript U) of PPKID 4-8, KID-AB, KID-B and peptide C were also determined (Table 1.7). As expected, KID-ABU and KID-BU peptides possess very low affinities for HisKIX (Figure 1.10). Only a small change in polarization of the KID-ABU-Flu (61 mP) or KID-BU-Flu (76 mP) molecules was observed even at the highest HisKIX concentrations tested (150 µM and 325 µM, respectively). This experiment allows us to place a lower limit on the Kd of the complex formed between each of these peptides and HisKIX. If we estimate the change in polarization of KID-ABU-Flu to be 110 mP and the change in polarization of KID-BU-Flu to be 150 mP when fully bound by HisKIX (based on observed changes in polarization of 116 mP for fully HisKIX-bound KID-ABp and 161 mP for KID-Bp), we can estimate that the Kd of the KID-ABU•HisKIX complex must be greater than 116 µM and the Kd of the KID-BU•HisKIX complex must be greater than 297 µM. Remarkably, the seven amino acid changes (including the conservative Tyr to Phe mutation) that convert KID-BU to peptide CU dramatically enhance CBP KIX-binding affinity (ΔΔG ≥ -1.5 kcal•mol⁻¹). Peptide CU binds HisKIX with a dissociation constant of 21.5 ± 2.6 µM. Addition of the turn and selected polyproline helix regions to yield peptides PPKIDU 6-8 increases CBP KIX-binding affinity. PPKID6U, PPKID7U and PPKID8U exhibit exceptionally high affinity for HisKIX, as measured by fluorescence polarization, with dissociation constants ranging from 1.5 µM to 3.1 µM (Figures 1.11). These values correspond to at least 37- to 77-fold enhancements in HisKIX-binding affinity compared to KID-ABU and at least 96- to 198-fold enhancements relative to KID-
Figure 1.10 Binding curves for unphosphorylated control peptides. The affinity of KID-AB$^{U}$, KID-B$^{U}$ and peptide C$^{U}$ for HisKIX were measured by equilibrium fluorescence polarization. The polarization of fluorescein-labeled peptide (peptide$^{U}$Flu) was monitored as a function of HisKIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID$^{U}$Flu bound values using the $P_{\text{min}}$ and $P_{\text{max}}$ values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID$^{U}$Flu bound values to equilibrium binding equation 1.2. The $K_{d}$’s derived from this analysis for the peptide$^{U}$Flu•HisKIX interaction under these conditions are greater than 116 µM (KID-AB$^{U}$), greater than 297 µM (KID-B$^{U}$) and 21.5 ± 2.6 µM (peptide C$^{U}$).
Figure 1.11 Binding curves for unphosphorylated PPKID peptides. The affinity of PPKID4\textsuperscript{U}, PPKID5\textsuperscript{U}, PPKID6\textsuperscript{U}, PPKID7\textsuperscript{U} and PPKID8\textsuperscript{U} for HisKIX were measured by equilibrium fluorescence polarization. The KID-AB\textsuperscript{U} binding curve is shown for comparison. The polarization of fluorescein-labeled peptide (peptide\textsuperscript{U}Flu) was monitored as a function of HisKIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID\textsuperscript{U}Flu bound values using the $P_{\text{min}}$ and $P_{\text{max}}$ values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID\textsuperscript{U}Flu bound values to equilibrium binding equation 1.2. The $K_d$’s derived from this analysis for the peptide\textsuperscript{U}Flu•HisKIX interaction under these conditions are $12.1 \pm 2.4 \mu$M (PPKID4\textsuperscript{U}), $6.6 \pm 2.0 \mu$M (PPKID5\textsuperscript{U}), $1.5 \pm 0.1 \mu$M (PPKID6\textsuperscript{U}), $2.3 \pm 0.2 \mu$M (PPKID7\textsuperscript{U}), $3.1 \pm 0.5 \mu$M (PPKID8\textsuperscript{U}) and greater than 116 $\mu$M (KID-AB\textsuperscript{U}).
Furthermore, peptides PPKIDU 6-8 bind HisKIX with 7- to 14-fold enhancements in binding affinity when compared to peptide CU. Thus, the selected polyproline helix and turn regions of the PPKIDU 6-8 peptides contribute -1.2 to -1.6 kcal•mol⁻¹ to the free energy of complex formation with KIX. PPKID4 and PPKID5, two peptides selected from both the phosphorylated and unphosphorylated library bind HisKIX worse than PPKID6U⁻8U with equilibrium dissociation constants of 12.1 ± 2.4 µM and 6.6 ± 2.0 µM.

The HisKIX-binding affinities of two variants of PPKID6 were also investigated. Each variant contained a simple modification of residue Ser18: phosphorylation and substitution of serine by glutamic acid. Surprisingly phosphorylation (PPKID6P) of this peptide yielded only a 2-fold improvement in binding KIX (Kd = 624 nM) over the unphosphorylated peptide (Figure 1.12). This is a significantly smaller difference than observed for phosphorylated and unphosphorylated versions of other selected peptides or KID-AB. The KIX-binding affinity of a serine to glutamic acid mutant of PPKID6U (PPKID6 S18E) was also measured. Glutamic acid can mimic a phosphoserine, however there are conflicting reports as to whether this substitution is accepted in the context of the KID-KIX interaction. One report shows that a peptide with glutamic acid in place of a serine leads to a binding affinity intermediate between phosphorylated and unphosphorylated serine in the same position. Other studies show no effect of the glutamic acid. In the context of PPKID6, replacement of the serine with glutamic acid yielded a peptide with lower affinity for KIX (Kd = 10.9 µM) than both phosphorylated and unphosphorylated versions of PPKID6. There are two possible explanations for this result. First of all, S133 is part of the N-cap of the α-helix in KID. Replacing this serine with a glutamic acid in the selected peptide may decrease the helical propensity of the
Figure 1.12 Binding curves for PPKID6 peptides. The affinity of PPKID6\textsuperscript{P}, PPKID6\textsuperscript{U} and PPKID6 S18E for HisKIX were measured by equilibrium fluorescence polarization. The polarization of fluorescein-labeled peptide (PPKID6Flu) was monitored as a function of HisKIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID6Flu bound values using the P\textsubscript{min} and P\textsubscript{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID6Flu bound values to equilibrium binding equation 1.2. The K\textsubscript{d}’s derived from this analysis for the PPKID6Flu•HisKIX interaction under these conditions are 624 ± 49 nM (PPKID6\textsuperscript{P}), 1.5 ± 0.1 µM (PPKID6\textsuperscript{U}) and 10.9 ± 2.0 µM (PPKID6 S18E).
peptide enough to significantly disrupt binding. Secondly, the selected peptide may not be binding KIX in a manner similar to KID. As a result, replacing the serine with a negatively charged residue creates an unfavorable interaction.

**Affinity of PPKID peptides for GST-KIX**

The phosphorylated KID domain (KID-AB^p) has a high affinity for GST-KIX; the complex was previously determined to have an equilibrium dissociation constant of 1.2 ± 0.5 μM, comparable to that measured with HisKIX (562 ± 41 nM). Likewise, the poor affinity of the unphosphorylated CREB domain (KID-AB^u) for GST-KIX and HisKIX are identical (greater than 120 μM^70 versus greater than 116 μM). The affinity of PPKID6^u for GST-KIX is comparable to its affinity for HisKIX (1.9 ± 0.2 μM versus 1.5 ± 0.1 μM) (Figure 1.13), therefore PPKID6^u does not appear to have an affinity for the GST-tag. However, many of the selected peptides bind to GST-KIX much better than HisKIX. For example PPKID4^p, the phosphorylated peptide with the highest affinity for HisKIX (515 ± 44 nM), binds GST-KIX with very high affinity (15 ± 1.9 nM).

**Competition of PPKID peptides with KID-AB^p**

Two sets of experiments were performed to investigate the binding modes of PPKID4^p and PPKID6^u. First, competition fluorescence polarization experiments assessed the ability of PPKID4^p and PPKID6^u to compete with KID-AB^p for binding CBP KIX. In particular, the fraction of fluorescently tagged PPKID4^p or PPKID6^u bound to HisKIX at equilibrium was monitored as a function of the concentration of unlabeled KID-AB^p. For competition experiments with PPKID4^p the concentration of HisKIX was kept constant at 1.5 μM, for competition experiments with PPKID6^u the concentration of

61
Figure 1.13 Binding curves for selected PPKID peptides with GST-KIX. The affinity of PPKID$^{4P}$ and PPKID$^{6U}$ for GST-KIX were measured by equilibrium fluorescence polarization. The polarization of fluorescein-labeled peptide (peptideFlu) was monitored as a function of GST-KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKIDFlu bound values using the $P_{\text{min}}$ and $P_{\text{max}}$ values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKIDFlu bound values to equilibrium binding equation 1.2. The $K_d$’s derived from this analysis for the peptideFlu•GST-KIX interaction under these conditions are 15 ± 1.9 nM (PPKID$^{4P}$) and 1.9 ± 0.2 µM (PPKID$^{6U}$).
Fraction of PPKID Bound

Flu

<table>
<thead>
<tr>
<th>Flu</th>
<th>PPKID4P</th>
<th>PPKID6U</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
</tr>
</tbody>
</table>

[GST-KIX] (M)

Fraction of PPKID\textsuperscript{Flu} Bound

- **PPKID4\textsuperscript{P}**
- **PPKID6\textsuperscript{U}**

![Graph showing the fraction of PPKID bound to GST-KIX at different concentrations.](image-url)
HisKIX was 3 µM (because PPKID6<sup>U</sup> has a slightly higher K<sub>d</sub>). The concentration of PPKID4<sup>p</sup>-flu or PPKID6<sup>U</sup>-flu was constant at 25 nM and the concentration of unlabeled KID-AB<sup>p</sup> was varied from 179 nM - 89 µM. HisKIX, PPKID-flu and KID-AB<sup>p</sup> were incubated at 25 °C for 60 minutes. Binding was measured by equilibrium fluorescence polarization. Raw polarization data was converted to fraction of PPKID bound and the data was fit to a mathematical equation for a sigmoidal curve. These experiments reveal that KID-AB<sup>p</sup> competes with both PPKID4<sup>p</sup> and PPKID6<sup>U</sup> for binding to CBP KIX (Figure 1.14). The concentration of KID-AB<sup>p</sup> needed to displace 50% of fluorescently tagged PPKID4<sup>p</sup> or PPKID6<sup>U</sup> from HisKIX (the IC<sub>50</sub> value) is 3.2 µM or 2.4 µM, respectively. Due to the conditions of the experiment, an IC<sub>50</sub> but not a K<sub>d</sub> could be determined because the concentration of the target protein HisKIX used is well above the K<sub>d</sub>. As expected given the conditions of the assay, these values are slightly higher than the K<sub>d</sub> of the KID-AB<sup>p</sup>•HisKIX complex determined by direct fluorescence polarization analysis (562 ± 41 nM). These results indicate that HisKIX cannot interact simultaneously with KID-AB<sup>p</sup> and either PPKID4<sup>p</sup> or PPKID6<sup>U</sup>, and are consistent with an interaction of both PPKID4<sup>p</sup> and PPKID6<sup>U</sup> within the CREB KID<sup>p</sup>-binding cleft of CBP KIX. Although KID-AB<sup>p</sup> competes with both PPKID4<sup>p</sup> and PPKID6<sup>U</sup> for binding to HisKIX, small changes in KID-AB<sup>p</sup> concentration around the corresponding IC<sub>50</sub> values have a larger effect on the change in the fraction of PPKID4<sup>p</sup> bound than on the change in the fraction of PPKID6<sup>U</sup> bound. This result suggests that differences may exist in the orientation or geometry of PPKID4<sup>p</sup> and PPKID6<sup>U</sup> when bound to CBP KIX.

Secondly, competition fluorescence polarization experiments were performed with flu-labeled PPKID4<sup>p</sup> and unlabeled PPKID4<sup>p</sup> to verify that the fluorescein label did
Figure 1.14 Competition between KID-AB\textsuperscript{p} and PPKID4\textsuperscript{p} or PPKID6\textsuperscript{U} for binding to HisKIX measured by fluorescence polarization. Serial dilutions of KID-AB\textsuperscript{p} were incubated with 1.5 µM or 3.0 µM HisKIX and 25 nM fluorescein-labeled PPKID4\textsuperscript{p} or PPKID6\textsuperscript{U} (peptide\textsuperscript{Flu}) for 60 min at 25 °C, respectively. Each point represents an average of three independent samples; the error bars denote standard error. Observed polarization values were converted to fraction of peptide\textsuperscript{Flu} bound using experimentally determined P\textsubscript{min} and P\textsubscript{max} values corresponding to the polarization of samples containing 25 nM peptide\textsuperscript{Flu} alone and peptide\textsuperscript{Flu} with 1.5 µM or 3.0 µM HisKIX, respectively. Curves shown represent the best fit of fraction of peptide\textsuperscript{Flu} bound values to equation (3). The close agreement between the K\textsubscript{d} of the KID-AB\textsuperscript{p}•HisKIX complex and the IC\textsubscript{50} values determined here provides evidence that the fluorescein tag appended to KID-AB\textsuperscript{p} contributes neither positively nor negatively to the stability of the KID-AB\textsuperscript{p}•HisKIX complex.
not contribute to binding. While a quantitative value could not be determined due to aggregation of PPKID4\textsuperscript{P} at high concentrations, I could qualitatively show that unlabeled PPKID4\textsuperscript{P} was competing with the labeled version. Competition experiments in the presence and absence of HisKIX identified the upper and lower limits of expected polarization values. At values above 5 μM, peptide aggregation interfered with the polarization, however the polarization does increase between 5 μM (pol=0.20) and 157 nM (pol=0.248) as expected if unlabeled PPKID4\textsuperscript{P} is competing for the same binding site (Figure 1.15).

**Structure of PPKID peptides**

Circular dichroism spectra of the selected phosphorylated and unphosphorylated peptides do not exhibit an α-helical signature in 10 mM potassium phosphate buffer at 4 °C, however with the addition of 30-50% 2,2,2-trifluoroethanol (TFE), the peptides display a significant amount of helicity (Figures 1.16-1.21). TFE is known to stabilize α-helices when there is intrinsic helical propensity.\textsuperscript{72} In our system, adding TFE to the solvent system creates a partly hydrophobic and aqueous environment, which may be a better mimic of the natural binding site of PPKID than aqueous solution as the PPKID are expected to bind to a hydrophobic patch on CBP KIX. The unphosphorylated peptide with the highest affinity for CBP KIX, PPKID6\textsuperscript{U} is 56% α-helical in 50% TFE, as determined by the mean residue ellipticity at 222 nm (Figure 1.16). This peptide demonstrates the highest α-helical content of all peptides tested. Unphosphorylated peptides PPKID7\textsuperscript{U} and PPKID8\textsuperscript{U} also exhibit increased α-helicity in the presence of TFE; 40% and 28% helical, respectively in 50% TFE (Figure 1.17 and Figure 1.18).
Figure 1.15 Competition between HisKIX-bound PPKID4\(^{\text{flu}}\) and unlabeled PPKID4\(^{P}\) monitored by fluorescence polarization. Error bars denote the standard deviation. As expected, the polarization of fluorescently labeled competitor decreases. Polarization of fluorescently labeled PPKID4\(^{P}\) in the presence of competitor (where 74% of PPKID4\(^{P}\) is bound by HisKIX) and in the absence of competitor and HisKIX (unbound PPKID4\(^{P}\)) are shown for reference. Although PPKID4\(^{P}\) aggregates in the absence of CBP KIX at concentrations above 5 \(\mu\)M, and the change in polarization resulting from this aggregation prevents us from calculating and IC\(_{50}\) value, the data at PPKID4\(^{P}\) concentrations of 5 \(\mu\)M and 2.5 \(\mu\)M clearly show that unlabeled PPKID4\(^{P}\) competes effectively with PPKID4\(^{P}\).
Polarization (mP)

- 157 nM
- 2.5 μM
- 5 μM
- unlabeled PPKID4P
- HisKIX

Polarization (mP)
**Figure 1.16** CD spectra of 10 µM PPKID6 in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
PPKID6**

Wavelength (nm)

MRE (mdeg cm$^2$ dmol$^{-1}$ res$^{-1}$)

0% TFE
15% TFE
30% TFE
50% TFE
**Figure 1.17** CD spectra of 10 μM PPKID7 in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
PPKID7^U

MRE (mdeg cm^2 dmol^{-1} res^{-1})

Wavelength (nm)

-15000 -10000 -5000 0 5000 10000 15000 205 215 225 235 245 255

0% TFE
30% TFE
50% TFE
**Figure 1.18** CD spectra of 10 μM PPKID8 in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
Wavelength (nm)

MRE (mdeg cm$^2$ dmol$^{-1}$ res$^{-1}$)
Figure 1.19 CD spectra of 10 μM PPKID4° in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
Wavelength (nm)

MRE (mdeg cm² dmol⁻¹ res⁻¹)

0% TFE
30% TFE
50% TFE

PPKID4°
Figure 1.20 CD spectra of 10 μM PPKID5° in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
Wavelength (nm)

MRE (mdeg cm$^2$ dmol$^{-1}$ res$^{-1}$)

PPKID5°

0% TFE
30% TFE
50% TFE
**Figure 1.21** CD spectra of 10 µM PPKID6° in 10 mM potassium phosphate (pH 7.0) with varying amounts of trifluoroethanol. The spectra were acquired at 4 °C from 195 nm to 260 nm with 1.0 nm step size and 2 second averaging at each point. Each spectrum shown is the mean of three scans and is background corrected and smoothed over three points.
Wavelength (nm)

MRE (mdeg cm$^2$ dmol$^{-1}$ res$^{-1}$)

-15000

-10000

-5000

0

50% TFE

30% TFE

0% TFE

PPKID6$^\circ$

195 205 215 225 235 245 255

Wavelength (nm)
Surprisingly, PPKID7\textsuperscript{U} exhibited similar helical content in both 30% and 50% TFE, whereas PPKID8\textsuperscript{U} only showed slight helicity in 30% TFE.

The helical structure of the selected phosphorylated peptides PPKID4\textsuperscript{P} and PPKID5\textsuperscript{P} and the phosphorylated version of PPKID6 were also investigated. All three peptides show increased \(\alpha\)-helicity in the presence of TFE. PPKID4\textsuperscript{P} and PPKID5\textsuperscript{P} are 37% and 38% helical in 50\% TFE (Figure 1.19 and Figure 1.20). PPKID6\textsuperscript{P} is 32\% helical, much lower than PPKID6\textsuperscript{U} (Figure 1.21). In addition, the shape of the CD spectra is noticeably different between phosphorylated and unphosphorylated peptides suggesting the phosphoserine may have an effect on TFE-induced \(\alpha\)-helicity.

**Affinity of PPKID peptides for CBP KIX variants**

PPKID4\textsuperscript{P} and PPKID6\textsuperscript{U} compete with KID-AB\textsuperscript{P} for binding to CBP KIX, therefore, a more detailed binding site for these ligands was investigated. A set of twelve well-studied CBP KIX variants\textsuperscript{62} were used to compare the binding location and orientation of PPKID4\textsuperscript{P} and PPKID6\textsuperscript{U} to that of KID-AB\textsuperscript{P}, whose interactions with CBP KIX have been characterized in detail by NMR.\textsuperscript{60} These twelve CBP KIX variants each contain a single alanine or phenylalanine substitution within or around the CREB-binding groove (Figure 1.22), and their affinities for KID-AB\textsuperscript{P} span a 3.7 kcal\textbullet mol\textsuperscript{-1} range. It can be reasoned, if PPKID4\textsuperscript{P} and PPKID6\textsuperscript{U} interact with CBP in a manner that mimics that of KID-AB\textsuperscript{P}, their affinities for these twelve variants should parallel those of KID-AB\textsuperscript{P}. The relative affinities of KID-AB\textsuperscript{P}, PPKID4\textsuperscript{P} and PPKID6\textsuperscript{U} for the panel of CBP KIX variants were measured by equilibrium fluorescence polarization (Table 1.8).
**Figure 1.22** Overview of CBP KIX variants. Residues on CBP KIX (gray surface) that are individually substituted with alanine are labeled and highlighted in blue. The CREB KID is shown in ribbon representation with the phosphoserine shown explicitly. Below is a list of interactions between the CBP KIX residues in the variant panel and KID-AB\textsuperscript{p} derived from the NMR structure.
## CBP KIX - KID-AB\(^{\scriptscriptstyle P}\) Interactions

<table>
<thead>
<tr>
<th>CBP KIX</th>
<th>KID-AB(^{\scriptscriptstyle P}) Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>L599</td>
<td>Hydrophobic P146, A145</td>
</tr>
<tr>
<td>L603</td>
<td>Hydrophobic A145, L141</td>
</tr>
<tr>
<td>K606</td>
<td>Hydrophobic L141</td>
</tr>
<tr>
<td>Y650</td>
<td>Hydrophobic A145, L141, L138</td>
</tr>
<tr>
<td>LL652-3</td>
<td>652-no contacts,</td>
</tr>
<tr>
<td></td>
<td>653-Hydrophobic L141</td>
</tr>
<tr>
<td>E655</td>
<td>Hydrophobic Y134</td>
</tr>
<tr>
<td>I657</td>
<td>Hydrophobic L141, L138, I137</td>
</tr>
<tr>
<td>Y658</td>
<td>H-bond with phosphate,</td>
</tr>
<tr>
<td></td>
<td>Hydrophobic Y134, L128</td>
</tr>
<tr>
<td>I660</td>
<td>No contacts</td>
</tr>
<tr>
<td>Q661</td>
<td>No contacts</td>
</tr>
<tr>
<td>K662</td>
<td>Salt bridge with phosphate</td>
</tr>
</tbody>
</table>
Table 1.8 Binding affinity of PPKID and CBP KIX variants

<table>
<thead>
<tr>
<th></th>
<th>K&lt;sub&gt;d&lt;/sub&gt;(µM)</th>
<th>ΔΔG (kcal•mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>K&lt;sub&gt;d&lt;/sub&gt;(µM)</th>
<th>ΔΔG (kcal•mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>K&lt;sub&gt;d&lt;/sub&gt;(µM)</th>
<th>ΔΔG (kcal•mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBP KIX</td>
<td>KID-AB&lt;sup&gt;P&lt;/sup&gt;</td>
<td></td>
<td>PPKID4&lt;sup&gt;P&lt;/sup&gt;</td>
<td></td>
<td>PPKID6&lt;sup&gt;U&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>wild-type</td>
<td>0.41± 0.04</td>
<td>0.61± 0.04</td>
<td>0.54± 0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoserine contacts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y658A</td>
<td>142± 12</td>
<td>3.5</td>
<td>3.9± 0.3</td>
<td>1.1</td>
<td>1.7± 0.4</td>
<td>0.68</td>
</tr>
<tr>
<td>Y658F</td>
<td>26± 5</td>
<td>2.5</td>
<td>4.1± 0.2</td>
<td>1.1</td>
<td>2.8± 0.4</td>
<td>0.97</td>
</tr>
<tr>
<td>K662A</td>
<td>4.8± 0.4</td>
<td>1.5</td>
<td>3.9± 0.3</td>
<td>1.1</td>
<td>1.8± 0.3</td>
<td>0.71</td>
</tr>
<tr>
<td>Hydrophobic contacts within cleft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L599A</td>
<td>1.1± 0.1</td>
<td>0.58</td>
<td>1.6± 0.1</td>
<td>0.58</td>
<td>3.3± 0.3</td>
<td>1.1</td>
</tr>
<tr>
<td>L603A</td>
<td>3.4± 0.3</td>
<td>1.2</td>
<td>3.4± 0.5</td>
<td>1.0</td>
<td>2.5± 0.3</td>
<td>0.91</td>
</tr>
<tr>
<td>K606A</td>
<td>2.3± 0.2</td>
<td>1.0</td>
<td>3.1± 0.2</td>
<td>0.97</td>
<td>1.2± 0.1</td>
<td>0.47</td>
</tr>
<tr>
<td>Y650A</td>
<td>9.4± 0.7</td>
<td>1.8</td>
<td>3.0± 0.3</td>
<td>0.95</td>
<td>1.3± 0.2</td>
<td>0.52</td>
</tr>
<tr>
<td>LL652-3AA</td>
<td>2.9± 0.2</td>
<td>1.2</td>
<td>2.9± 0.2</td>
<td>0.93</td>
<td>0.14± 0.02</td>
<td>-0.80</td>
</tr>
<tr>
<td>I657A</td>
<td>1.9± 0.1</td>
<td>0.9</td>
<td>2.7± 0.1</td>
<td>0.89</td>
<td>3.6± 0.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Hydrophobic contacts outside cleft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E655A</td>
<td>0.71± 0.05</td>
<td>0.32</td>
<td>0.54± 0.05</td>
<td>-0.07</td>
<td>1.7± 0.2</td>
<td>0.68</td>
</tr>
<tr>
<td>I660A</td>
<td>0.27± 0.03</td>
<td>-0.26</td>
<td>1.1± 0.1</td>
<td>0.35</td>
<td>6.1± 0.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Q661A</td>
<td>0.93± 0.07</td>
<td>0.48</td>
<td>1.1± 0.1</td>
<td>0.35</td>
<td>0.41± 0.06</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

Hydrophobic contacts outside cleft:
- E655A: 0.71± 0.05, ΔΔG = -0.07 kcal•mol<sup>-1</sup>, ΔΔG = 1.7 kcal•mol<sup>-1</sup>, ΔΔG = 0.68 kcal•mol<sup>-1</sup>
- I660A: 0.27± 0.03, ΔΔG = -0.26 kcal•mol<sup>-1</sup>, ΔΔG = 6.1 kcal•mol<sup>-1</sup>, ΔΔG = 1.4 kcal•mol<sup>-1</sup>
- Q661A: 0.93± 0.07, ΔΔG = 0.48 kcal•mol<sup>-1</sup>, ΔΔG = 1.1 kcal•mol<sup>-1</sup>, ΔΔG = 0.41 kcal•mol<sup>-1</sup>
Comparison of PPKID4\textsuperscript{p} and KID-AB\textsuperscript{p} - phosphoserine contacts

Recognition of the phosphoserine residue in KID-AB\textsuperscript{p} by Y658 and K662 in CBP KIX contributes significantly to the stability of the KID\textbullet KIX complex, with energetic contributions between 1.5 and 4 kcal\textbullet mol\textsuperscript{-1}.\textsuperscript{62,66,68} The Y658 side chain donates a phenolic hydrogen bond to one terminal phosphoserine oxygen whereas the K662 ammonium group forms a salt bridge with a second oxygen (Figure 1.23).\textsuperscript{60,62} Y658F and K662A CBP KIX both exhibit significantly decreased affinity for KID-AB\textsuperscript{p}, consistent with previous results,\textsuperscript{62} with equilibrium dissociation constants of 26 ± 5 and 4.8 ± 0.4 \textmu M, respectively (Figure 1.24). These values correspond to less favorable binding free energies (2.5 and 1.5 kcal\textbullet mol\textsuperscript{-1}) than the wild type KID-AB\textsuperscript{p}\textbullet CBP KIX complex. The free energy changes measured with these two variants, as well as the Y658A variant (discussed below), are consistent with previous work and available structural information.\textsuperscript{60,62}

Y658F and K662A CBP KIX also exhibit significantly decreased affinity for PPKID4\textsuperscript{p} (Figure 1.25). The equilibrium dissociation constant of the PPKID4\textsuperscript{p}\textbullet Y658F complex is 4.1 ± 0.2 \textmu M. This value corresponds to a binding free energy that is 1.1 kcal\textbullet mol\textsuperscript{-1} less favorable than that of the wild type PPKID4\textsuperscript{p}\textbullet CBP KIX complex, approximately one half the magnitude of the change seen with KID-AB\textsuperscript{p}. Comparably, the equilibrium dissociation constant of the PPKID4\textsuperscript{p}\textbullet K662A complex is 3.9 ± 0.3 \textmu M; this value corresponds to a binding free energy that is 1.1 kcal\textbullet mol\textsuperscript{-1} less favorable than that of the wild type PPKID4\textsuperscript{p}\textbullet CBP KIX complex. These data support a model where Y658 and the K662 each contribute significantly to the affinity of PPKID4\textsuperscript{p} for CBP KIX. Interestingly, residue Y658, which forms a hydrogen bond with KID-AB\textsuperscript{p}, is more
Figure 1.23 Close-up view of phosphoserine contacts in the KID-AB<sup>P</sup>•CBP KIX complex. The backbones of KID-AB<sup>P</sup> and CBP KIX are shown as red and blue ribbons, respectively; the side chains of Y658, K662 (from CBP KIX) and S133 (from KID-AB<sup>P</sup>) are shown explicitly.
Figure 1.24 Binding curves for KID-AB\textsuperscript{p} with alanine variants of CBP KIX phosphate contacts. The affinity of KID-AB\textsuperscript{p} for Y658A, Y658F and K662A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (KID-AB\textsuperscript{p}\textbullet CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (KID-AB\textsuperscript{p}Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of KID-AB\textsuperscript{p}Flu bound values using the \( P_{\text{min}} \) and \( P_{\text{max}} \) values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of KID-AB\textsuperscript{p}Flu bound values to equilibrium binding equation 1.2. The \( K_d \)'s derived from this analysis for the KID-AB\textsuperscript{p}Flu\textbullet CBP KIX interaction under these conditions are 410 ± 40 nM (CBP KIX), 142 ± 12 µM (Y658A), 26 ± 5 µM (Y658F), and 4.8 ± 0.4 µM (K662A).
Fraction of KID-AB\textsuperscript{P} Bound

<table>
<thead>
<tr>
<th>Protein</th>
<th>Y658A</th>
<th>Y658F</th>
<th>K662A</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBP KIX (KID-AB\textsuperscript{P})</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

[y-axis label: Fraction of KID-AB\textsuperscript{P} Bound]

[x-axis label: [KIX] (M)]
Figure 1.25 Binding curves for PPKID4\textsuperscript{p} with alanine variants of CBP KIX phosphate contacts. The affinity of PPKID4\textsuperscript{p} for Y658A, Y658F and K662A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID4\textsuperscript{p}\textbullet CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID4\textsuperscript{p}Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID4\textsuperscript{p}Flu bound values using the P_{min} and P_{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID4\textsuperscript{p}Flu bound values to equilibrium binding equation 1.2. The K_d’s derived from this analysis for the PPKID4\textsuperscript{p}Flu\textbullet CBP KIX interaction under these conditions are 610 ± 40 nM (CBP KIX), 3.9 ± 0.3 µM (Y658A), 4.1 ± 0.2 µM (Y658F), and 3.9 ± 0.3 µM (K662A).
Fraction of PPKID4 Bound

P Flu

[KIX] (M)

Y658A
Y658F
K662A

CBP KIX (PPKID4^P)

Fraction of PPKID4^P Flu Bound

[10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}]

0 0.2 0.4 0.6 0.8 1
important overall for KID-AB$^p$ than for PPKID4$^p$, whereas the potential salt bridge-forming residue K662 has a similar contribution for both peptides.

The affinities of KID-AB$^p$ and PPKID4$^p$ for the Y658A variant of CBP KIX in which the entire tyrosine side chain has been replaced by alanine were also examined. The NMR structure of the KID-AB$^p•$CBP KIX complex shows this aromatic ring packed against residue L128 on helix A of KID-AB$^p$. This variant binds KID-AB$^p$ with exceptionally low affinity ($K_d = 142 \pm 12 \mu M$), a loss in binding free energy of 3.5 kcal•mol$^{-1}$. Presumably this large effect is observed because the complex suffers from loss of both the phenolic hydrogen bond and an important hydrophobic packing interaction. Since PPKID4$^p$ lacks a residue corresponding to L128 on helix A, one would expect the stability of the PPKID4$^p•$Y658A complex to be comparable to that of the PPKID4$^p•$Y658F complex. Indeed, the equilibrium dissociation constant of the PPKID4$^p•$Y658A complex is $4.1 \pm 0.2 \mu M$, corresponding to a free energy loss of 1.1 kcal•mol$^{-1}$, a value identical to that measured for the Y658F•PPKID4$^p$ complex.

*Comparison of PPKID4$^p$ and KID-AB$^p$- hydrophobic contacts*

Next, those residues that line the shallow KID-AB$^p$ binding cleft on CBP KIX were considered. Six of the twelve variants (L599A, L603A, K606A, Y650A, LL652-3AA and I657A) contain alanine in place of a residue within this cleft. For example, Y650 of CBP KIX comprises one face of the binding cleft and interacts with three hydrophobic side chains of KID-AB$^p$, including L138, L141 and A145 on KID-AB$^p$ (Figure 1.26). The Y650A variant binds KID-AB$^p$ with significantly reduced affinity ($K_d = 9.4 \pm 0.7 \mu M$), corresponding to a 1.8 kcal•mol$^{-1}$ loss in binding energy (Figure 1.27). Together, residues L603 and K606 form one side of the binding cleft of CBP KIX,
Figure 1.26 Close-up view of hydrophobic contacts in the KID-AB$^p$•CBP KIX complex. The backbones of KID-AB$^p$ and CBP KIX are shown as red and blue ribbons, respectively; the side chains of L603, K606 and Y650 (from CBP KIX) and L138, L141 and A145 (from KID-AB$^p$) are shown explicitly.
Figure 1.27 Binding curves for KID-AB$^p$ with alanine variants of CBP KIX hydrophobic contacts within the binding cleft. The affinity of KID-AB$^p$ for L599A, L603A, K606A, Y650A, LL652-3AA and I657A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (KID-AB$^p$•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (KID-AB$^p$Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of KID-AB$^p$Flu bound values using the $P_{\min}$ and $P_{\max}$ values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of KID-AB$^p$Flu bound values to equilibrium binding equation 1.2. The $K_d$’s derived from this analysis for the KID-AB$^p$Flu•CBP KIX interaction under these conditions are $410 \pm 40$ nM (CBP KIX), $1.1 \pm 0.1$ µM (L599A), $3.4 \pm 0.3$ µM (L603A), $2.3 \pm 0.2$ µM (K606A), $9.4 \pm 0.7$ µM (Y650A), $2.9 \pm 0.2$ µM (LL652-3AA), and $1.9 \pm 0.1$ µM (I657A).
Fraction of KID-AB\textsuperscript{P} Bound

<table>
<thead>
<tr>
<th></th>
<th>Flu</th>
<th>CBP KIX</th>
<th>L599A</th>
<th>L603A</th>
<th>K606A</th>
<th>Y650A</th>
<th>LL652-3AA</th>
<th>I657A</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^{-8}</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>10^{-7}</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>10^{-6}</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>10^{-5}</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>10^{-4}</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
which interact with CREB residues L141 and A145. The L603A and K606A variants bind KID-AB\textsuperscript{p} with equilibrium dissociation constants of 3.4 ± 0.3 and 2.3 ± 0.2 µM, corresponding to losses in binding free energy of 1.2 and 1.0 kcal\textbullet{}mol\textsuperscript{-1} compared to wild type. Other CBP KIX residues that comprise part of the hydrophobic cleft are L653 and I657; both interact with CREB residue L141 in addition to other residues. As expected, variants LL652-3AA and I657A also bind KID-AB\textsuperscript{p} with lower affinity ($K_d$ = 2.9 ± 0.2 and 1.9 ± 0.1 µM), corresponding to losses in binding free energy of 1.2 and 0.9 kcal\textbullet{}mol\textsuperscript{-1}, respectively. L599 interacts with only one residue, P146 from CREB, at the bottom of CBP’s binding cleft. The L599A variant binds KID-AB\textsuperscript{p} with lower affinity, but to a lesser extent than other variants that make up the hydrophobic cleft; the KID-AB\textsuperscript{p}•L599A complex has an equilibrium dissociation constant of 1.1 ± 0.1 µM, corresponding to a loss in binding free energy of 0.58 kcal\textbullet{}mol\textsuperscript{-1}. Thus, as expected, all CBP KIX variants of residues that make hydrophobic contacts with CREB have diminished affinity for KID-AB\textsuperscript{p}.

All six CBP KIX variants within the hydrophobic binding cleft also show diminished affinity for PPKID4\textsuperscript{p}, ($K_d$’s ranging from 1.6 ± 0.1 to 3.4 ± 0.5 µM) (Figure 1.28). These dissociation constants correspond to the free energy changes between 0.58 and 1.0 kcal\textbullet{}mol\textsuperscript{-1}. With one exception, these values are in close agreement with the range observed among complexes with KID-AB\textsuperscript{p} ($\Delta\Delta G = 0.58$ to 1.8 kcal\textbullet{}mol\textsuperscript{-1}). The one exception is Y650A CBP KIX, whose complex with KID-AB\textsuperscript{p} is 1.8 kcal\textbullet{}mol\textsuperscript{-1} less stable than the complex with wild type CBP KIX. By contrast, the PPKID4\textsuperscript{p}•Y650A complex is only 0.95 kcal\textbullet{}mol\textsuperscript{-1} less stable than the complex with wild type CBP KIX.
**Figure 1.28** Binding curves for PPKID4\(^p\) with alanine variants of CBP KIX hydrophobic contacts within the binding cleft. The affinity of PPKID4\(^p\) for L599A, L603A, K606A, Y650A, LL652-3AA and I657A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID4\(^p\)•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID4\(^p\)Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID4\(^p\)Flu bound values using the \(P_{\text{min}}\) and \(P_{\text{max}}\) values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID4\(^p\)Flu bound values to equilibrium binding equation 1.2. The \(K_d\)’s derived from this analysis for the PPKID4\(^p\)Flu•CBP KIX interaction under these conditions are 610 ± 40 nM (CBP KIX), 1.6 ± 0.1 μM (L599A), 3.4 ± 0.5 μM (L603A), 3.1 ± 0.2 μM (K606A), 3.0 ± 0.3 μM (Y650A), 2.9 ± 0.2 μM (LL652-3AA), and 2.7 ± 0.1 μM (I657A).
Fraction of PPKID4 Bound

<table>
<thead>
<tr>
<th>Protein</th>
<th>KIX (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L599A</td>
<td>$10^{-8}$</td>
</tr>
<tr>
<td>L603A</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>K606A</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>Y650A</td>
<td>$10^{-5}$</td>
</tr>
<tr>
<td>LL652-3AA</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>I657A</td>
<td>$10^{-3}$</td>
</tr>
</tbody>
</table>

**Graph:**
- CBP KIX (PPKID4*)
- L599A
- L603A
- K606A
- Y650A
- LL652-3AA
- I657A

**Axes:**
- X-axis: [KIX] (M)
- Y-axis: Fraction of PPKID4 Bound

**Legend:**
- CBP KIX (PPKID4*)
- L599A
- L603A
- K606A
- Y650A
- LL652-3AA
- I657A
Comparison of PPKID4<sup>P</sup> and KID-AB<sup>P</sup>- residues surrounding the binding pocket

In addition to the variants described above which contain mutations within the KID-AB<sup>P</sup> binding pocket, three variants – E655A, I660A and Q661A – that surround the binding pocket were also examined. These variants may provide additional information about the subtle differences in binding of these ligands. The equilibrium dissociation constants of the complexes of these variants with KID-AB<sup>P</sup> fall between 0.27 ± 0.03 and 0.93 ± 0.07 µM (Figure 1.29), values very close to that of the wild type complex (ΔΔG = −0.26 and +0.48 kcal•mol<sup>−1</sup>, respectively). These three CBP KIX variants bind PPKID4<sup>P</sup> with equilibrium dissociation constants between 0.54 ± 0.05 and 1.1 ± 0.1 µM (Figure 1.30), corresponding to free energy changes between -0.07 and 0.35 kcal•mol<sup>−1</sup>; similar to KID-AB<sup>P</sup>, these values are also close to those of the wild type complex.

Comparison of PPKID6<sup>U</sup> and KID-AB<sup>P</sup>- phosphoserine contacts

PPKID6<sup>U</sup> lacks the phosphoserine that dominates the energetics of the PPKID4<sup>P</sup>•CBP KIX and KID-AB<sup>P</sup>•CBP KIX complexes. Therefore, if this ligand is bound in a manner that mirrors that of PPKID4<sup>P</sup> and KID-AB<sup>P</sup>, it should be less sensitive to changes at positions that contact the phosphoserine, Y658 and K662 of CBP KIX. Surprisingly, PPKID6<sup>U</sup> binds both variants with significantly decreased affinity relative to wild type CBP KIX (Figure 1.31). The equilibrium dissociation constants of the Y658F•PPKID6<sup>U</sup> and K662A•PPKID6<sup>U</sup> complexes are 2.8 ± 0.4 and 1.8 ± 0.3 µM, corresponding to free energy losses of 0.97 and 0.71 kcal•mol<sup>−1</sup>, respectively, relative to the wild type complex. Interestingly, the Y658A variant binds PPKID6<sup>U</sup> better (K<sub>d</sub> = 1.7 ± 0.4 µM) than the Y658F variant (K<sub>d</sub> = 2.8 ± 0.4 µM), whereas Y658A and Y658F bind
Figure 1.29 Binding curves for KID-AB° with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft. The affinity of KID-AB° for E655A, I660A and Q661A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (KID-AB°•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (KID-AB°Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of KID-AB°Flu bound values using the P_{min} and P_{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of KID-AB°Flu bound values to equilibrium binding equation 1.2. The K_{d}'s derived from this analysis for the KID-AB°Flu•CBP KIX interaction under these conditions are 410 ± 40 nM (CBP KIX), 0.71 ± 0.05 µM (E655A), 0.27 ± 0.03 µM (I660A), and 0.93 ± 0.07 µM (Q661A).
Figure 1.30 Binding curves for PPKID4<sup>P</sup> with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft. The affinity of PPKID4<sup>P</sup> for E655A, I660A and Q661A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID4<sup>P</sup>•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID4<sup>P</sup>Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID4<sup>P</sup>Flu bound values using the P<sub>min</sub> and P<sub>max</sub> values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID4<sup>P</sup>Flu bound values to equilibrium binding equation 1.2. The K<sub>d</sub>’s derived from this analysis for the PPKID4<sup>P</sup>Flu•CBP KIX interaction under these conditions are 610 ± 40 nM (CBP KIX), 0.54 ± 0.05 µM (E655A), 1.1 ± 0.1 µM (I660A), and 1.1 ± 0.1 µM (Q661A).
Fraction of PPKID4 $^\text{P}$ Bound

$[\text{KIX}]$ (M)

- I660A
- E655A
- CBP KIX (PPKID4$^\text{P}$)
- Q661A

Fraction of PPKID4$^\text{P}$ Flu Bound

[KIX] (M)
Figure 1.31 Binding curves for PPKID6\textsuperscript{U} with alanine variants of CBP KIX phosphate contacts. The affinity of PPKID6\textsuperscript{U} for Y658A, Y658F and K662A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID6\textsuperscript{U}•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID6\textsuperscript{U}Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID6\textsuperscript{U}Flu bound values using the P\textsubscript{min} and P\textsubscript{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID6\textsuperscript{U}Flu bound values to equilibrium binding equation 1.2. The K\textsubscript{d}'s derived from this analysis for the PPKID6\textsuperscript{U}Flu•CBP KIX interaction under these conditions are 540 ± 60 nM (CBP KIX), 1.7 ± 0.4 μM (Y658A), 2.8 ± 0.4 μM (Y658F), and 1.8 ± 0.3 μM (K662A).
Fraction of PPKID6 Bound

U Flu

[KIX] (M)

Y658A
Y658F
K662A

CBP KIX
(PPKID6*)

10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}

0
0.2
0.4
0.6
0.8
1

Fraction of PPKID6* Flu Bound

[KIX] (M)
PPKID4\textsuperscript{p} equally well. These results are not consistent with a model in which the \( \alpha \)-helix in PPKID6\textsuperscript{U} is positioned within the CBP KIX cleft in a manner that closely mimics that of KID-AB\textsuperscript{p}. In contrast, they support a model characterized by an overlapping, but alternate, binding mode for PPKID6\textsuperscript{U} compared to PPKID4\textsuperscript{p} and KID-AB\textsuperscript{p}.

**Comparison of PPKID6\textsuperscript{U} and KID-AB\textsuperscript{p}-hydrophobic contacts**

To further explore differences in the binding modes of PPKID4\textsuperscript{p} and PPKID6\textsuperscript{U}, the affinity of PPKID6\textsuperscript{U} for the six CBP KIX variants with substitutions of KIX side chains that line the KID-AB\textsuperscript{p} binding pocket were measured. Substitution of these KIX side chains (variants L599A, L603A, K606A, Y650A and I657A) resulted in complexes with PPKID6\textsuperscript{U} that are 0.47 to 1.1 kcal\( \text{mol}^{-1} \) less stable than the wild type complex (Figure 1.32). In contrast, the complex with LL652-3AA is 0.8 kcal\( \text{mol}^{-1} \) more stable. Variants L599A, L603A, and I657A exhibit moderately diminished binding affinity to PPKID6\textsuperscript{U} (\( K_d = 3.3 \pm 0.3, 2.5 \pm 0.3, \) and \( 3.6 \pm 0.4 \) \( \mu \)M; \( \Delta \Delta G = 1.1, 0.91, \) and 1.1 kcal\( \text{mol}^{-1} \), respectively), whereas K606A and Y650A show smaller decreases in PPKID6\textsuperscript{U} binding affinity (\( K_d = 1.2 \pm 0.1 \) and \( 1.3 \pm 0.2 \) \( \mu \)M; \( \Delta \Delta G = 0.47 \) and 0.52 kcal\( \text{mol}^{-1} \), respectively). Analysis of the differences in affinity among this set of PPKID6\textsuperscript{U}•CBP KIX complexes reveals a pattern very different from that observed for the analogous KID-AB\textsuperscript{p} or PPKID4\textsuperscript{p} complexes. For example, Y650 and L599 make the largest and smallest energetic contributions, respectively, to binding of KID-AB\textsuperscript{p}, whereas Y650 contributes least and L599 contributes most to complex formation with PPKID6\textsuperscript{U}. 
**Figure 1.32** Binding curves for PPKID6_U with alanine variants of CBP KIX hydrophobic contacts within the binding cleft. The affinity of PPKID6_U for L599A, L603A, K606A, Y650A, LL652-3AA and I657A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID6_U•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID6_U•Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID6_U•Flu bound values using the P_min and P_max values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID6_U•Flu bound values to equilibrium binding equation 1.2. The K_d’s derived from this analysis for the PPKID6_U•Flu•CBP KIX interaction under these conditions are 540 ± 60 nM (CBP KIX), 3.3 ± 0.3 µM (L599A), 2.5 ± 0.3 µM (L603A), 1.2 ± 0.1 µM (K606A), 1.3 ± 0.2 µM (Y650A), 0.14 ± 0.02 µM (LL652-3AA), and 3.6 ± 0.4 µM (I657A).
Fraction of PPKID6 bound

U Flu

\[ [KIX] (M) \]

L599A
L603A
K606A
Y650A
LL652-3AA
I657A

CBP KIX
(PPKID6^U)

10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}

Fraction of PPKID6\textsuperscript{U} Flu Bound

[KIX] (M)
Comparison of PPKID6U and KID-ABp-residues surrounding the binding pocket

Although CBP KIX variants E655A, I660A and Q661A display affinities for KID-ABp and PPKID4p comparable to wild type CBP KIX, two of these variants show significantly diminished affinity for PPKID6U (Figure 1.33). Variant I660A exhibits the largest decrease in PPKID6U binding affinity over all CBP KIX variants in the panel, with an equilibrium dissociation constant of $6.1 \pm 0.7 \mu M$, corresponding to a free energy loss of $1.4 \text{ kcal} \cdot \text{mol}^{-1}$. E655A also exhibits decreased affinity for PPKID6U, albeit to a lesser extent ($K_d = 1.7 \pm 0.2 \mu M; \Delta \Delta G = 0.68 \text{ kcal} \cdot \text{mol}^{-1}$), whereas Q661A binds PPKID6U with affinity comparable to wild type CBP KIX. Again, differences in the relative and absolute contributions of CBP KIX residues to the energy of complex formation with PPKID6U compared to PPKID4p and KID-ABp suggest an alternate, but overlapping binding site for PPKID6U.

Affinity of PPKID4p variants for CBP KIX

The results described above suggest that the $\alpha$-helix in PPKID4p is positioned within the CBP KIX cleft in a manner that closely mimics that of KID-ABp. While PPKID4p displays only nascent $\alpha$-helicity in water in the absence of CBP KIX, as determined by circular dichroism, it exhibits intrinsic $\alpha$-helicity as demonstrated by the increase in helicity in TFE. However, this data sheds no light on whether PPKID4p is bound in a folded aPP-like hairpin conformation or an open conformation where the polyproline helix makes additional contacts on CBP KIX.

If PPKID4p folds upon binding within the CBP KIX binding cleft into an aPP-like conformation, then the stability of its complex with CBP KIX should depend on the
Figure 1.33 Binding curves for PPKID6\textsuperscript{U} with alanine variants of CBP KIX hydrophobic contacts outside of the binding cleft. The affinity of PPKID6\textsuperscript{U} for E655A, I660A and Q661A CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID6\textsuperscript{U}•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID6\textsuperscript{U}Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID6\textsuperscript{U}Flu bound values using the P\textsubscript{min} and P\textsubscript{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID6\textsuperscript{U}Flu bound values to equilibrium binding equation 1.2. The $K_d$’s derived from this analysis for the PPKID6\textsuperscript{U}Flu•CBP KIX interaction under these conditions are $540 \pm 60$ nM (CBP KIX), $1.7 \pm 0.2$ µM (E655A), $6.1 \pm 0.7$ µM (I660A), and $0.41 \pm 0.06$ µM (Q661A).
presence of residues that comprise the intact hydrophobic core and the corresponding variants would display diminished affinity for CBP KIX. However, if the N-terminal region of PPKID4\(^p\) interacts with the CBP KIX surface elsewhere, then the stability of the PPKID4\(^p\)•CBP KIX complex would not depend on the identities of putative folding residues from the PPKID4\(^p\) \(\alpha\)-helix because these residues would be solvent-exposed. In this case, alanine variants would result in affinity for CBP KIX comparable to wild type PPKID4\(^p\). To test this hypothesis, a set of eleven PPKID4\(^p\) variants were prepared, in which alanine or sarcosine is substituted for a PPKID4\(^p\) residue that comprises the hydrophobic core in the putative folded state (Figure 1.34).\(^{49}\) These variants include alanine substitutions along the face of the PPKID4\(^p\) \(\alpha\)-helix opposite the face used to contact CBP KIX (L17A, F20A, L24A, Y27A and L28A) and six variants with alanine or sarcosine substitutions along the PPII helix (P2A, P2Z, P5A, P5Z, P8A and P8Z).

Fluorescence polarization analysis was used to determine the CBP KIX binding affinities of these eleven PPKID4\(^p\) variants in which one residue within the aPP hydrophobic core had been substituted with alanine. The equilibrium dissociation constants of the PPKID4\(^p\) variant•CBP KIX complexes range from 0.68 ± 0.05 to 3.09 ± 0.18 \(\mu\)M, corresponding to binding energies between 0.07 and 0.96 kcal\(\cdot\)mol\(^{-1}\) less favorable than the wild type complex (Figure 1.35, Figure 1.36 and Table 1.9). The stabilities of the variant complexes fall into three categories. The least stable complexes containing variants F20A and Y27A were 0.85 and 0.96 kcal\(\cdot\)mol\(^{-1}\) less stable than the wild type complex; moderately stable complexes containing variants P5A, P8A and L24A were 0.3 to 0.38 kcal\(\cdot\)mol\(^{-1}\) less stable than the wild type complex. Six of the variants, P2A, P2Z, P5Z, P8Z, L17A and L28A, formed CBP KIX complexes with
Figure 1.34 Side chain packing in the αP hydrophobic core. This view illustrates the relative orientation of F20, L24, Y27, and L28 on the α-helix (blue) and P3 and P5 on the PPII helix (yellow). Residues shown in blue and yellow were individually substituted with alanine or sarcosine in PPKID4 and their binding to CBP KIX was measured.
Figure 1.35 Binding curves for PPKID4\(^\alpha\)-helix variants with CBP KIX. The affinity of L17A, F20A, L24A, Y27A and L28A PPKID4\(^\alpha\) for CBP KIX were measured by equilibrium fluorescence polarization. The binding curve for the wild-type interaction (PPKID4\(^\alpha\)•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID4\(^\alpha\)Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID4\(^\alpha\)Flu bound values using the \( P_{\text{min}} \) and \( P_{\text{max}} \) values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID4\(^\alpha\)Flu bound values to equilibrium binding equation 1.2. The \( K_d \)'s derived from this analysis for the PPKID4\(^\alpha\)Flu•CBP KIX interaction under these conditions are 610 ± 40 nM (CBP KIX), 0.68 ± 0.05 µM (L17A), 2.55 ± 0.25 µM (F20A), 1.16 ± 0.07 µM (L24A), 3.09 ± 0.18 µM (Y27A) and 0.83 ± 0.11 µM (L28A).
Figure 1.36 Binding curves for PPKID4\textsuperscript{p} proline variants with CBP KIX. The affinity of P2A, P2Z, P5A, P5Z, P8A and P8Z PPKID4\textsuperscript{p} for CBP KIX were measured by equilibrium fluorescence polarization; Z = sarcosine variant. The binding curve for the wild-type interaction (PPKID4\textsuperscript{p}•CBP KIX) is shown for reference. The polarization of fluorescein-labeled peptide (PPKID4\textsuperscript{p}Flu) was monitored as a function of CBP KIX concentration after incubation of the binding reaction for 30 min at room temperature. Each point represents the average of three independent samples and the error bars denote the standard error. Polarization values were converted to fraction of PPKID4\textsuperscript{p}Flu bound values using the P\textsubscript{min} and P\textsubscript{max} values derived from the best fit of the polarization data to equilibrium binding equation 1.1. The curve shown is the best fit of the fraction of PPKID4\textsuperscript{p}Flu bound values to equilibrium binding equation 1.2. The K\textsubscript{d}'s derived from this analysis for the PPKID4\textsuperscript{p}Flu•CBP KIX interaction under these conditions are 610 ± 40 nM (CBP KIX), 0.87 ± 0.04 μM (P2A), 0.83 ± 0.08 μM (P2Z), 1.02 ± 0.09 μM (P5A), 0.80 ± 0.03 μM (P5Z), 1.07 ± 0.08 μM (P8A) and 0.78 ± 0.05 μM (P8Z).
Fraction of Ligand Bound

Flu

P2A

P2Z

P5A

P5Z

P8A

P8Z

10^{-7} 10^{-6} 10^{-5}

KIX (M)

PPKID4p

P2A

P2Z

P5A

P5Z

P8A

P8Z

Fraction of Ligand Bound

KIX (M)
Table 1.9  CBP KIX-binding affinity of PPKID4P variants

<table>
<thead>
<tr>
<th>PPKID4P</th>
<th>$K_d$ (µM)</th>
<th>$\Delta \Delta G$ (kcal·mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>0.61 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>Polyproline helix variants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2A</td>
<td>0.87 ± 0.04</td>
<td>0.21</td>
</tr>
<tr>
<td>P2Z</td>
<td>0.83 ± 0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>P5A</td>
<td>1.02 ± 0.09</td>
<td>0.30</td>
</tr>
<tr>
<td>P5Z</td>
<td>0.80 ± 0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>P8A</td>
<td>1.07 ± 0.08</td>
<td>0.33</td>
</tr>
<tr>
<td>P8Z</td>
<td>0.78 ± 0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>$\alpha$-helix variants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L17A</td>
<td>0.68 ± 0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>F20A</td>
<td>2.55 ± 0.25</td>
<td>0.85</td>
</tr>
<tr>
<td>L24A</td>
<td>1.16 ± 0.07</td>
<td>0.38</td>
</tr>
<tr>
<td>Y27A</td>
<td>3.09 ± 0.18</td>
<td>0.96</td>
</tr>
<tr>
<td>L28A</td>
<td>0.83 ± 0.11</td>
<td>0.18</td>
</tr>
</tbody>
</table>
stabilities that were similar (0.07 to 0.21 kcal\(\text{mol}^{-1}\)) to the wild type complex. Interestingly, those side chains that contribute to the stability of the PPKID4\(^\text{p}\)•CBP KIX complex – F20, L24, Y27, P5 and P8 – all lie at the center of the aPP hydrophobic core (Figure 1.34). In the structure of aPP,\(^{49}\) the aromatic side chain of F20 inserts between the side chains of residues P3 and P5, the branched side chain of L24 packs against P8 and F20, and the side chain of Y27 packs against P8. By contrast, those side chains that contribute minimally to stability – P2, L17 and L28 – lie at the edge of the hydrophobic core of aPP and participate in fewer van der Waals interactions. Thus, these results support a model in which PPKID4\(^\text{p}\) folds (although to a lesser extent than aPP) upon binding to CBP KIX into an aPP-like hairpin conformation.

**Discussion**

Protein grafting is an adaptable strategy that has been used to identify miniature protein ligands for the DNA major groove,\(^{50-52,73}\) deep protein clefts,\(^{53}\) and certain PPII recognition proteins.\(^{55}\) With the identification of the PPKID ligands for CBP KIX, this approach was extended to the recognition of shallow protein surfaces, and for the first time a post-translational modification was incorporated into the selection protocol. In addition, the binding mode and orientation of PPKID4\(^\text{p}\) and PPKID6\(^\text{U}\) were investigated in detail to determine if these ligands target the same site on the CBP KIX surface as the natural ligand CREB KID, on which they were based.
Phosphorylated miniature protein ligands for CBP KIX

Two peptides, PPKID4 and PPKID5, were identified through a phosphorylated phage display selection. Despite sharing only two of five residues in the randomized region (GPSQPTYP versus GLSWPTYH), these two peptides bind the CBP KIX domain with identical affinity when phosphorylated (515 nM and 534 nM). PPKID4\(^\text{P}\) and PPKID5\(^\text{P}\) display CBP KIX-binding affinities comparable to that of the full length phosphorylated CREB KID domain (KID-AB\(^\text{P}\)) and 100-fold greater than that of phosphorylated helix B of CREB KID (KID-B\(^\text{P}\)), which corresponds to the portion of CREB KID that was grafted onto the aPP scaffold. Thus, in the context of phosphorylated peptides, the aPP scaffold and selected residues enhance the CBP KIX-binding affinity of the CREB KID helix B functional epitope comparable to the 92-fold enhancement provided by the addition of helix A. In addition, the increases in affinity seen with phosphorylated PPKID (2.7 kcal•mol\(^{-1}\)) is comparable to values observed in previous applications of protein grafting in which the miniature protein target contains a much deeper binding pocket.\(^{53}\)

Interestingly, peptide C\(^\text{p}\), which contains the residues comprising the putative \(\alpha\)-helical portion of the PPKID peptides, has significantly higher affinity for CBP KIX (\(K_d = 2.4 \mu\text{M}\)) than does KID-B\(^\text{P}\) (\(K_d = 51.6 \mu\text{M}\)). Peptides C\(^\text{p}\) and KID-B\(^\text{P}\) differ at seven of twenty residues, including the conservative Tyr to Phe mutation, which is known to have no effect on binding in the context of the full-length CREB KID domain.\(^{63}\) The enhanced CBP KIX-binding affinity exhibited by peptide C\(^\text{p}\) relative to KID-B\(^\text{P}\) may result from direct contacts between the aPP-derived residues in peptide C\(^\text{p}\) and the CBP KIX surface, or by virtue of increased helical propensity of aPP-derived residues in
peptide C\(^p\) compared to the corresponding residues in KID-B\(^p\), or by a combination of these two factors. The 1.8 kcal\(\cdot\)mol\(^{-1}\) increase in stability of the peptide C\(^p\)•KIX complex compared to the KID-B\(^p\)•KIX complex is significantly larger than the 0.9 kcal\(\cdot\)mol\(^{-1}\) increase in stability of the PPKID\(^p\)•KIX complexes compared to the peptide C\(^p\)•KIX complex. Thus, in the context of phosphorylated peptides, changes within the \(\alpha\)-helical region contribute more to CBP KIX-binding affinity than residues (even selected residues) within the polyproline helix region. Despite the lack of conservation among the residues of the putative polyproline helix regions of PPKID\(^p\) and PPKID5\(^p\), it seems likely that these molecules bind to CBP KIX in a similar manner, as most of their CBP KIX-binding affinity is derived from their common \(\alpha\)-helical portion.

Although the PPKID library members contain all the residues from CREB KID helix B known to be important for KID-AB\(^p\)•KIX complex formation, no residues from helix A of CREB KID were included in the library. This might be expected to limit the affinity of the PPKID library since at least one residue from helix A, L128, makes energetically significant contacts in the KID-AB\(^p\)•KIX complex; when L128 is mutated to alanine, CBP KIX binding is abolished in a GST pull-down assay.\(^62\) The structure of the KID-AB\(^p\)•KIX complex indicates that L128 packs against Y658.\(^60\) This observation is confirmed by binding data where KID-AB\(^p\) exhibits weaker affinity (1.0 kcal\(\cdot\)mol\(^{-1}\)) for KIX variant Y658A, than Y658F due to removal of a hydrophobic packing interaction. Since the PPKID libraries omit a L128 homologue, it is not surprising that PPKID\(^p\) binds equally to Y658A and Y658F. The identification of the PPKID peptides, which have high affinity for CBP KIX despite the absence of residues from helix A, represents an example of a successful protein grafting application using a thermodynamically
incomplete set of binding residues. This data reinforces findings of another protein grafting experiment, the recent example of a miniature engrailed homeodomain which binds DNA with high affinity despite the absence of residues known to be critical for formation of the natural engrailed\*DNA complex.\(^2\)

**Comparison to unphosphorylated counterparts**

Peptides PPKID4 and PPKID5 were also selected as unphosphorylated ligands and exhibit moderate affinity for CBP KIX. Phosphorylation of PPKID4 and PPKID5 increases binding to CBP KIX 23- and 12-fold, respectively. This is in contrast to KID-AB, which requires phosphorylation to achieve high affinity for CBP KIX, exhibiting a greater than 206-fold enhancement (\(\Delta \Delta G > 3.2 \text{ kcal} \cdot \text{mol}^{-1}\)) in CBP KIX-binding affinity upon phosphorylation. Peptide C, like the PPKID peptides, exhibits less dependence on phosphorylation for high affinity CBP KIX binding; upon phosphorylation, the affinity of peptide C for CBP KIX is enhanced 9-fold. These results suggest that the relatively high affinity of the unphosphorylated PPKID peptides and peptide C\(^U\) is derived from their common \(\alpha\)-helical region. As is true in the context of phosphorylated peptides, peptide C\(^U\) exhibits significantly increased affinity for CBP KIX when compared to KID-B\(^U\) (\(\Delta \Delta G > -1.6 \text{ kcal} \cdot \text{mol}^{-1}\)). However, peptide C\(^U\) is also a better ligand for CBP KIX than the unphosphorylated full-length CREB KID domain, KID-AB\(^U\) (\(\Delta \Delta G > -1.0 \text{ kcal} \cdot \text{mol}^{-1}\)). Thus, the seven amino acid changes which convert KID-B\(^U\) to peptide C\(^U\) provide a significant increase in CBP KIX-binding affinity, whereas the addition of helix A to KID-B\(^U\) does not result in high affinity CBP KIX binding (\(K_d > 116 \mu \text{M}\)). In the context of phosphorylated peptides, the seven residue changes which convert KID-B\(^P\) to peptide C\(^P\)
provide an enhancement in CBP KIX-binding affinity comparable to the addition of helix A to KID-B\textsuperscript{p}.

This difference illustrates the different factors that play a role in binding between phosphorylated and unphosphorylated ligands. Phosphorylated ligands rely heavily on phosphate interactions, which contribute to half of the binding free energy of the KID-AB\textsuperscript{p}•KIX complex, the other half comes from hydrophobic interactions.\textsuperscript{62} On the other hand, unphosphorylated ligands rely entirely on hydrophobic interactions. All seven amino acid changes, which convert KID-B\textsuperscript{U} to peptide C\textsuperscript{U}, substitute hydrophobic residues - three leucines, two phenylalanines and two tyrosines – replacing only one hydrophobic residue from KID-B\textsuperscript{U}. With this large increase in hydrophobicity it is not surprising that peptide C\textsuperscript{U} exhibits high affinity for the hydrophobic groove on CBP KIX.

\textit{Unphosphorylated miniature protein ligands for CBP KIX}

Despite the difficulties with the GST-KIX unphosphorylated selection, the consensus reached at the randomized positions for the two unphosphorylated selections are remarkably similar: Leu or Ile at position 2, Trp at position 4, Pro at position 5 and aromatic or negatively charged residues at position 7. Not surprisingly, one of the peptides from the GST-KIX selection (PPKID7) was also selected in the second selection when GST-KIX and HisKIX were alternated as the target protein, indicating the validity of this selection method.

The five PPKID peptides identified in unphosphorylated selections (PPKID\textsuperscript{U} 4-8) display a wider range of affinities for CBP KIX than is observed among the phosphorylated PPKID peptides, with $K_d$s from 1.5 $\mu$M to 12.1 $\mu$M. Notably, the three peptides identified solely in the selections lacking a phosphorylation step (PPKID\textsuperscript{U} 6-8)
exhibit the highest affinity for CBP KIX, with $K_d$s from 1.5 $\mu$M to 3.1 $\mu$M. This result provides further evidence that the library members in phosphorylated selections were successfully phosphorylated by PKA, as the PPKID molecules identified in those selections (PPKID 4-5) display significantly lower CBP KIX-binding affinity when unphosphorylated than PPKID$^U$ 6-8.

The best unphosphorylated peptide, PPKID6$^U$, binds CBP KIX with extremely high affinity ($K_d = 1.5 \mu$M), only two-fold worse than the full length phosphorylated CREB KID domain, and at least 200-fold better than unphosphorylated helix B of CREB KID ($K_d > 297 \mu$M). The aPP scaffold contributes at least 3.1 kcal$\cdot$mol$^{-1}$ stability to the PPKID6$^U$•KIX complex, of which 1.5 kcal$\cdot$mol$^{-1}$ is due to the putative polyproline helix and turn region (including the selected residues). The dependence of the CBP KIX-binding affinity of PPKID6 on phosphorylation ($\Delta\Delta G = -0.5$ kcal$\cdot$mol$^{-1}$) was significantly smaller than is observed for PPKID 1-5 ($\Delta\Delta G = -1.0$ to -2.2 kcal$\cdot$mol$^{-1}$).\textsuperscript{70} Notably, PPKID6$^U$ binds CBP KIX with 8-fold greater affinity than PPKID4$^U$, which contains all aPP-derived residues at the randomized positions. In the unphosphorylated case, then, the simple graft of the functional epitope of CREB KID helix B onto the aPP scaffold did not produce the optimal CBP KIX-binding molecule and a functional selection successfully identified a molecule with significantly higher affinity.

Interestingly, despite these observed differences in the binding modes of PPKID6$^U$ and KID-AB$^p$, PPKID6$^U$ and KID-AB$^p$ cannot bind simultaneously to CBP KIX, as KID-AB$^p$ and PPKID6$^U$ compete for binding HisKIX in an equilibrium fluorescence polarization assay. These results are consistent with overlapping, but not identical, binding sites of PPKID6$^U$ and KID-AB$^p$ on CBP KIX.
**PPKID specificity**

In previous work, Dr. Stacey Rutledge investigated the ability of PPKID$_4^P$ and PPKID$_6^U$ to selectively bind CBP KIX. The specificity of PPKID$_4^P$ and PPKID$_6^U$ was determined by measuring their affinity for two globular proteins, carbonic anhydrase II and calmodulin, which are known to recognize hydrophobic or helical molecules. To determine the contributions arising from the selected polyproline helix and turn on the specificity of PPKID$_4^P$ and PPKID$_6^U$, the specificity of peptide C in both its phosphorylated and unphosphorylated forms was also examined.

The PPKID peptides bind carbonic anhydrase with low affinity, with $K_d$'s of 106 ± 12 µM and 79 ± 13 µM for PPKID$_4^P$ and PPKID$_6^U$, respectively. PPKID$_4^P$ preferentially binds HisKIX over carbonic anhydrase ($K_{rel} = 205$), which is considerably higher than the preference of control peptide C$_P$ for HisKIX over carbonic anhydrase ($K_{rel} = 40$). Similar conclusions are drawn when comparing PPKID$_6^U$ and peptide C$_U$; although these two molecules display similar affinities for carbonic anhydrase, with $K_d$ values of 79 µM and 66 µM, respectively, PPKID$_6^U$'s HisKIX preference ($K_{rel} = 53$) is significantly higher than peptide C$_U$'s ($K_{rel} = 3$). PPKID$_4^P$ and PPKID$_6^U$ also display a dramatic preference for binding CBP KIX over calmodulin. PPKID$_4^P$ binds calmodulin with a $K_d$ of 51 ± 12 µM, which corresponds to a $K_{rel}$ of 100. Peptide C$_P$ displays slightly lower specificity ($K_{rel} = 74$) than PPKID$_4^P$ for CBP KIX over calmodulin, a result of 5-fold lower affinity for HisKIX and 4-fold lower affinity for calmodulin ($K_d = 178 ± 42$ µM). The $K_d$ for the PPKID$_6^U$•calmodulin complex could not be determined definitively, but a lower limit of 168 µM could be placed on the $K_d$ value. Thus, PPKID$_6^U$, like
PPKID4\textsuperscript{P}, exhibits a significant preference for CBP KIX over calmodulin, with a specificity ratio of at least 112.

**Structure of PPKID**

Surprisingly, given their high affinity for CBP KIX, the PPKID peptides display only nascent $\alpha$-helicity in the absence of CBP KIX as determined using circular dichroism. Phosphorylation of the PPKID peptides has no effect on their $\alpha$-helical content, suggesting that the increased affinity of the phosphorylated PPKID peptides is not due to phosphorylation-dependent changes in their structures. However, the unphosphorylated peptides appear to have more intrinsic $\alpha$-helicity as illustrated by their TFE-induced structure. PPKID6\textsuperscript{P} is 56% helical in 50% TFE, whereas PPKID6\textsuperscript{U} is only 32% helical. In addition to the lower helical content for the phosphorylated version of PPKID6, the CD spectra of all phosphorylated peptides exhibit a different shape compared to the unphosphorylated peptides. This difference may indicate the phosphoserine has a different effect on TFE-induced $\alpha$-helicity. Other studies have shown the CREB KID domain is mostly disordered in the absence of CBP KIX, with less than 10% helicity observed in helix B.\textsuperscript{67} Furthermore, CD and NMR experiments have shown that phosphorylation of Ser133 has little or no effect on the folding of helix B,\textsuperscript{74,75} despite the effect of phosphorylation on CBP KIX-binding affinity. Based on mutagenesis of helix B of CREB KID, it has been proposed that the stability of CBP KIX-bound helix B rather than the stability of unbound helix B is the major determinant of CBP KIX-binding affinity.\textsuperscript{68} Interestingly, residues Arg135 and Lys136 of CREB KID have been suggested to contain a negative determinant to helicity.\textsuperscript{68,76} Two of the seven amino acid differences between peptide C and helix B correspond to these residues, and
this may, in part, explain the enhanced affinity of peptide C for CBP KIX and the high affinity of the unphosphorylated PPKID peptides for CBP KIX.

The lack of helical content in PPKID4, which corresponds to the simple graft of the CBP KIX-binding epitope onto the aPP scaffold, indicates that the changes made in the putative α-helical region are sufficient to disrupt the hydrophobic core of aPP, presumably due to the absence of Val30 and Val31 or the destabilizing effect of RRXS sequence (residues 15-18). We chose to randomize five residues from the polyproline helix which form part of the hydrophobic core of aPP, based on our success using a similar strategy to evolve DNA-binding miniature proteins. In theory it should be possible to identify members of the PPKID Library 1 that exhibit enhanced affinity due to stabilization of the CBP KIX-bound α-helical region by the randomized polyproline helix. In retrospect, this strategy may not have been the best one for this system, given the structure of the complex formed between KID-AB\(^p\) and CBP KIX. Although most of the energetically important contacts between KID-AB\(^p\) and CBP KIX are mediated by helix B, which binds into a shallow groove on the surface of CBP KIX, helix A contacts another region of the CBP KIX surface, and thereby contributes to CBP KIX-binding affinity. The conditions used in our functional selections would not distinguish between those library members that have enhanced affinity for CBP KIX by virtue of stabilization of the CBP KIX-bound α-helix and those that have enhanced affinity due to the gain of contacts between residues in the randomized region and the CBP KIX surface. Therefore, one explanation for the high affinity of the miniature proteins for CBP KIX despite their lack of well-defined structure is that the randomized region is acting akin to helix A, contacting the CBP KIX surface outside of the helix B-binding groove. It might
contact the same surface which binds helix A, or another region of the CBP KIX surface; at least two other proteins (c-Jun<sup>77</sup> and MLL<sup>78</sup>) bind to CBP KIX in ternary complexes with CREB KID, contacting distinct regions of the CBP KIX surface.

The minimal unphosphorylated activation domain of the proto-oncoprotein c-Myb (residues 291-315) binds constitutively to the KID-binding groove of the CBP KIX domain with a $K_d$ of 15 mM.<sup>76</sup> The c-Myb activation domain displays significant helical content (30%) in aqueous buffer and is 90% helical in the presence of 40% TFE.<sup>68</sup> In contrast to CREB KID, binding of c-Myb to CBP KIX is driven both by entropy and enthalpy.<sup>76</sup> Furthermore, $\alpha$-helix-destabilizing mutations of c-Myb significantly compromise CBP KIX binding. Taken together, these studies strongly suggest that the significant helicity of c-Myb is crucial for its high (as compared to unphosphorylated CREB KID) affinity recognition of CBP KIX. In light of these reports, it is quite surprising that the unphosphorylated PPKID peptides display affinities for CBP KIX up to 10-fold greater than the CBP KIX-binding affinity of c-Myb, since they lack of helical structure in the absence of CBP KIX.

**Binding mode and orientation of CBP ligands: PPKID<sup>4</sup>**

Two lines of evidence suggest that PPKID<sup>4</sup> interacts with the KID-AB<sup>P</sup>-binding groove on the CBP KIX surface in a manner similar to the natural ligand KID-AB<sup>P</sup>. First, PPKID<sup>4</sup> and KID-AB<sup>P</sup> compete for binding CBP KIX in an equilibrium fluorescence polarization assay. The concentration of KID-AB<sup>P</sup> needed to displace 50% of fluorescently tagged PPKID<sup>4</sup> is consistent with the $K_d$ of the KID-AB<sup>P</sup>*CBP KIX complex determined by direct fluorescence polarization analysis. Second, the stabilities of the PPKID<sup>4</sup> and KID-AB<sup>P</sup> complexes with CBP KIX variants verify that similar
residues on CBP KIX are important for binding of both ligands. These results confirm that protein grafting can produce ligands capable of high affinity recognition of a shallow hydrophobic groove such as found in CBP KIX.

Our results support a model where KID-AB\textsuperscript{p} and PPKID4\textsuperscript{p} bind in a similar manner to CBP KIX. This conclusion is best illustrated by Figure 1.37, in which the differences in stability (as determined by ΔΔG values in Table 1.8) between the KID-AB\textsuperscript{p} and PPKID4\textsuperscript{p} complexes of the CBP KIX variant panel are mapped on the structure of CBP KIX. In both cases, CBP KIX K662, which contacts the phosphoserine residue contributes equally to stability. Likewise, the CBP KIX residues that line the hydrophobic cleft, contacting CREB L141 (L603, K606, Y650, L653 and I657) are sensitive to alanine substitution leading to free energy changes in the complexes with KID-AB\textsuperscript{p} and PPKID4\textsuperscript{p} between 0.9-1.6 kcal\textsuperscript{-1}mol (residues colored yellow in Figure 1.37). The one exception is Y650 in the context of the KID-AB\textsuperscript{p}•Y650A complex, which contributes more to stability. Similarly, CBP KIX residues that contribute little (L599 and Q661; residues colored in green) to the binding of KID-AB\textsuperscript{p} also contribute little to the binding of PPKID4\textsuperscript{p}. The observation that the free energies of the PPKID4\textsuperscript{p}•CBP KIX complexes mirror those of the KID-AB\textsuperscript{p}•CBP KIX complexes provide strong evidence that the CBP KIX residues in question contribute in a similar way to complex formation with the two ligands.

For the most part, the binding free energies of CBP KIX variants for PPKID4\textsuperscript{p} mirror those for KID-AB\textsuperscript{p}, confirming that the CBP KIX residues in question make
Figure 1.37 Relative stabilities of the CBP KIX complexes of KID-AB\(^p\), PPKID4\(^p\) and PPKID6\(^U\) mapped on the structure of CBP KIX. Using the \(\Delta\Delta G\) values listed in Table 1.8, residues are colored according to the extent of which binding is affected compared to the complex with wild type KIX. Alanine variants forming complexes resulting in \(\Delta\Delta G\) values between 2.5-3.5 kcal\(\cdot\)mol\(^{-1}\) are shown in red, \(\Delta\Delta G = 1.7-2.4\) kcal\(\cdot\)mol\(^{-1}\) (orange), \(\Delta\Delta G = 0.9-1.6\) kcal\(\cdot\)mol\(^{-1}\) (yellow), \(\Delta\Delta G = 0.3-0.8\) kcal\(\cdot\)mol\(^{-1}\) (green), and \(\Delta\Delta G < 0.3\) kcal\(\cdot\)mol\(^{-1}\) (blue).
ΔΔG values (kcal·mol⁻¹)

- 2.5 - 3.5
- 1.7 - 2.4
- 0.9 - 1.6
- 0.3 - 0.8
- < 0.3

KID-AB⁺
PPKID₄⁺
PPKID₆⁻
similar energetic contributions to complex formation with the two ligands. However, two residues of CBP KIX (Y658, which makes a phosphoserine contact and Y650, which makes hydrophobic contacts) make considerably larger energetic contributions to complex formation with KID-AB\textsuperscript{p} than to complex formation with PPKID4\textsuperscript{p}. In the context of KID-AB\textsuperscript{p} binding, variants Y658F and Y650A decrease binding by 2.5 and 1.8 kcal\textsuperscript{-1} mol\textsuperscript{-1}, whereas with PPKID4\textsuperscript{p} these variants decrease binding by 1.1 and 0.95 kcal\textsuperscript{-1} mol\textsuperscript{-1}, respectively. Despite the diminished contributions of these two residues, CBP KIX displays approximately equal binding affinity for the two ligands. Taken together, these results suggest that some additional factor, whether enthalpic or entropic, contributes to the stability of the PPKID4\textsuperscript{p}•CBP KIX complex but not to that of the KID-AB\textsuperscript{p}•CBP KIX complex.

*Binding mode and orientation of CBP ligands: PPKID6\textsuperscript{U}*

PPKID6\textsuperscript{U} has a binding footprint on CBP KIX, which is distinct from, yet overlaps, that of either PPKID4\textsuperscript{p} or KID-AB\textsuperscript{p}. As shown in Figure 1.37, several residues (K606, Y650 and L653) which line the hydrophobic cleft surrounding CREB L141 make significantly smaller energetic contributions to binding PPKID6\textsuperscript{U} than they do for KID-AB\textsuperscript{p}. On the other hand, several side chains that contribute to KID-AB\textsuperscript{p}•CBP KIX stability, such as L603, I657 and Y658, also contribute significantly to PPKID6\textsuperscript{U}•CBP KIX stability. Interestingly, the CBP KIX residue that is the most energetically important for PPKID6\textsuperscript{U} binding is residue I660, which is on the opposite face of KIX from the KID-AB\textsuperscript{p}-binding site, and does not contribute to the stability of the wild-type complex. The residues key to recognition of PPKID6\textsuperscript{U} are widely scattered on the surface of CBP KIX,
suggesting that PPKID6\textsuperscript{U} may bind to CBP KIX by wrapping around the protein, making contacts with both the \(\alpha\)- and PPII-helix.

One reason why PPKID6\textsuperscript{U} is not directed into the CREB-binding pocket of CBP KIX in a manner similar to KID-AB\textsuperscript{P} may be the lack of a phosphoserine residue in this ligand. Yet surprisingly, variants of phosphoserine-contacting residues (Y658 and K662) moderately decrease PPKID6\textsuperscript{U} binding despite the absence of a phosphoserine residue. Unexpectedly, the CBP KIX variant Y658F disrupts PPKID6\textsuperscript{U} binding more than variant Y658A. This pattern of binding indicates that deletion of the tyrosine hydroxyl group is detrimental, whereas further deletion of the aromatic ring is actually beneficial in terms of PPKID6\textsuperscript{U} binding. Clearly, alternative interactions must be occurring between PPKID6\textsuperscript{U} and this region of the CBP KIX surface which are not found in the KID-AB\textsuperscript{P}•CBP KIX interface. These results, taken with competition results suggest that PPKID6\textsuperscript{U} binds an overlapping, but distinct, region of CBP KIX compared to CREB KID.

\textit{PPKID4\textsuperscript{P} Folding}

Results from the CBP KIX variant panel suggest the existence of one or more factors that contribute to the stability of the PPKID4\textsuperscript{P}•CBP KIX complex but not the KID-AB\textsuperscript{P}•CBP KIX complex. The data suggest that one factor could be the free energy gained as PKID4\textsuperscript{P} folds upon binding KIX. The relative affinities of the twelve PPKID4\textsuperscript{P} variants for KIX indicate that a small core of residues located on the (presumably) interior face of the PPKID4\textsuperscript{P} \(\alpha\)-helix (F20, L24 and Y27) are essential for binding to KIX. Substitution of each of these residues with alanine led to complexes with KIX that were 0.38 to 0.96 kcal\textsuperscript{-1}mol\textsuperscript{-1} less stable than the wild-type complex. In addition,
substitution of two residues located on the (presumably) interior face of the PPII helix (P5 and P8) with alanine led to complexes with KIX that were 0.30 to 0.33 kcal•mol⁻¹ less stable than the PPKID4⁺KIX complex. Notably, substitution of P5 or P8 with alanine eliminates the portion of the pyrolidine ring that would pack against α-helix residues F20 and Y27 (Figure 1.34). By contrast, substitution of P5 or P8 with sarcosine (P5Z and P8Z) retains the portion of the pyrolidine ring that would pack against α-helix residues F20 and Y27, allowing for hydrophobic packing. Indeed, the P5Z•KIX and P8Z•KIX complexes are more stable than the P5A•KIX and P8A•KIX complexes. These data suggest that PPKID4⁺ folds into an aPP-like conformation upon binding to KIX with residues P5, P8, F20, L24 and Y27 forming part of the hydrophobic core, and that this factor contributes to the stability of the PPKID4⁺•KIX complex.

Overall, alanine and sarcosine mutants of the three proline folding residues of PPKID4⁺ have relatively small effects on CBP KIX binding. The small effect of these residues on binding is not surprising considering the fact that a control peptide containing only the chimeric α-helix of PPKID peptides binds CBP KIX reasonably well when phosphorylated (Kᵋ = 2.4 μM), despite lacking the polyproline helix and β-turn.

**Experimental**

**General**

**Materials**

Complete protease inhibitor tablets were purchased from Roche Applied Science (Indianapolis, IN). Biotinylated thrombin, steptavidin agarose and Rosetta (DE3) heat
shock competent \textit{E. coli} cells were purchased from Novagen (Madison, WI). Electrocompetent XL1Blue \textit{E. coli} cells were purchased from Stratagene (La Jolla, CA). Glutathione sepharose, Nap-10 columns and M13K07 helper phage were purchased from Amersham Biosciences (Piscataway, NJ). Glutathione coated plates were purchased from Pierce Biotechnology (Rockford IL). Ni-NTA His-sorb microtiter 8-well strips were purchased from Qiagen (Valencia, CA). Corning Costar 384-well plates for fluorescence were purchased from Fisher Scientific (Pittsburgh, PA). 5-iodoacetamidofluorescein was purchased from Molecular Probes (Eugene, OR). Protein Kinase A (PKA) was purchased from Promega (Madison, WI). HisKIX (KIX with an N-terminal hexahistidine tag) was a gift from Stacey Rutledge.\textsuperscript{64} Reagents for peptide synthesis including rink amide resin and Fmoc-amino acids were purchased from Novabiochem (San Diego, CA) and solvents used with the Symphony Peptide Synthesizer were purchased from American Bioanalytical (Natick, MA).

\textit{Media and Buffers}

2X-YT media contains 16 g tryptone, 10 g yeast extract and 5 g NaCl per 1 L. LB media contains 10 g tryptone, 5 g yeast extract, 10 g NaCl and 1 ml 1 M NaOH per 1 L. SOB-AG agar contains 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 15 g Agar, 20 g glucose, 10 ml of MgCl\textsubscript{2}, and 100 mg ampicillin per 1 L.

KIX storage buffer contains 50 mM Tris (pH 8), 100 mM KCl, 12.5 mM MgCl\textsubscript{2}, 1 mM EDTA, 1 mM DTT and 0.05% Tween-20. Lysis buffer contains 50 mM potassium phosphate (pH 7.2), 150 mM NaCl, 1 mM DTT, 1% Tween-20 and a Complete protease inhibitor tablet (Roche). Elution buffer for glutathione sepharose contains 10 mM
glutathione and 50 mM Tris (pH 8). Thrombin cleavage buffer contains 20 mM Tris-HCl (pH 8), 150 mM NaCl, 2.5 mM CaCl₂, 1 mM DTT, and 0.05% Tween-20. PBST buffer contains 50 mM potassium phosphate (pH 7.2), 150 mM NaCl and 0.05% Tween-20. TBST buffer contains 20 mM Tris (pH 8.0), 150 mM NaCl and 0.05% Tween-20.

**Expression and Purification of GST-KIX and KIX variants**

pGST-ΔKIX(588-683) (a gift from Jennifer Nyborg) and pGST-KIX variants (a gift from Marc Montminy) were transformed into BL21(DE3)pArg electroporation competent or Rosetta (DE3) heat-shock competent *E. coli* cells. A single colony was used to inoculate a 1 L culture of LB media containing 0.2 mg/mL ampicillin with 0.05 mg/mL kanamycin (pArg cells) or 0.03 mg/ml chloramphenicol (Rosetta cells). The culture was incubated at 37 °C with shaking at 250 rpm until the solution reached an optical density of 0.6 absorbance units at 600 nm. Isopropyl β-D-thiogalactoside (IPTG) was added to a final concentration of 1 mM and incubation continued for 3 h at 37 °C. Cells were harvested by centrifugation for 20 min at 6,000 g and resuspended in 15 mL lysis buffer. Cells were lysed by sonication, insoluble material was pelleted by centrifugation for 30 min at 16,500 g, and the supernatant was retained. The supernatant was incubated with 1 mL glutathione sepharose overnight at 4 °C with shaking.

GST KIX – Column was packed by centrifugation at 1,000 g then washed twice with lysis buffer. Sepharose was incubated with 2 mL of elution buffer at 4 °C for 2 hours. After collecting the first eluent of GST-KIX, an additional 2 mL of elution buffer was incubated with the sepharose overnight. The two elutions were combined and desalted on a Nap-10 column. Pure protein was stored in KIX storage buffer.
KIX and KIX variants only - Glutathione sepharose was washed with 5 x 5 mL thrombin cleavage buffer. 10 units of biotinylated thrombin (Novagen) in 1 mL of thrombin cleavage buffer was added to glutathione sepharose and incubated with rotation at 25 °C for 4 h. Flow-through containing KIX protein was collected by centrifugation for 5 min. Biotinylated thrombin was removed from the protein sample by incubation with 100 µL streptavidin agarose for 30 min followed by centrifugation of sample in a spin filter at 500 g for 5 min. Proteins were desalted on a Nap-10 column and stored in KIX storage buffer. Protein concentration of wild-type KIX was determined with amino acid analysis, concentration of the KIX mutants were determined with Bio-Rad Detergent Compatible Protein Assay (based on the Lowry method), using wild-type KIX as a standard.

**Phage Display Selections**

The phage display library PPKID Library 1 was designed and constructed by Stacey Rutledge. A 1.5 mL glycerol stock of the initial pool (round 1) or output from the previous round (rounds 2-9) was used to inoculate 10 mL 2X YT-AG media (2X YT media containing 0.1 mg/mL ampicillin and 20 mg/mL glucose). The culture was incubated at 37 °C until it reached an optical density of 0.6 absorbance units at 600 nm. The culture was then infected with 4 x 10^{11} pfu M13K07 helper phage and incubated at 37 °C for 1 h. Cells were pelleted by centrifugation at 1,500 g, resuspended in 10 mL 2X YT-AK (2X YT media containing 0.1 mg/mL ampicillin and 0.05 mg/mL kanamycin) and incubated for 12-13 h at 37 °C. Cells were then pelleted by centrifugation at 1,500 g and the retained supernatant was filtered through a 0.45 µm syringe filter. Phage were precipitated with 1/5 volume PEG/NaCl (20% (w/v) PEG-8000, 2.5 M NaCl) on ice for
45 min, and then pelleted by centrifugation for 35 min at 24,000 g. For phosphorylated selections, the precipitated phage were resuspended in water and approximately 10^10 phage were phosphorylated in vitro with 2500 U PKA in 100 μM ATP, 40 mM Tris (pH 8) and 20 mM magnesium acetate for 2 h at 30 °C. Phosphorylated phage were precipitated on ice for 45 min with PEG/NaCl and then pelleted by centrifugation at maximum speed in a microcentrifuge for 30 min at 4 °C. Mock phosphorylation reactions were performed in parallel without PKA, and purified in the same manner. Precipitated phage (+/- PKA treatment) were resuspended in binding buffer for use in selections. Binding buffer was PBST buffer for HisKIX and TBST buffer for GST-KIX.

Selections against HisKIX were performed in Ni-NTA HisSorb microtiter 8-well strips and selections against GST-KIX were performed in glutathione-coated 96-well microtiter plates. 200 μL target protein was added to each well (the final concentration was 30 nM for GST-KIX and 100 nM for HisKIX) and incubated overnight with shaking at 4 °C. Wells were washed three times with HisKIX or GST-KIX binding buffer to remove unbound protein. For blocking, binding buffer containing 6% milk was added to each well and incubated at 4 °C for 3 h. After blocking, wells were washed three times with binding buffer. Phage purified as described were added to each well and incubated for 3 h at 4 °C or 25 °C. Nonbinding or weakly binding phage were removed by repeated washing (10 to 20 times, 1 min to 5 min in length, according to round) with binding buffer, see Table 1.5. Bound phage were eluted by incubation with 0.1 M glycine (pH 2.2) for 20 min. After neutralization of the eluted phage solution with 2 M Tris (pH 9.2), XL1 Blue E. coli cells in log phase were infected with input and output phage and incubated at 37 °C for 1 h. Serial dilutions of infected cells were plated on SOB-AG agar
plates. Cells infected with output phage were used to make glycerol stocks for subsequent rounds and stored at -70 °C.

**Peptide Sequences for Control Peptides**

Peptide sequences for the control peptides KID-AB and KID-B were based on the natural ligand CREB KID (residues 119-148 and 130-148, respectively). Peptide C is based on the α-helix of the PPKID Library 1. All peptides including PPKID peptides were synthesized with a glycine–valine linker (derived from CREB residues 149-150) and a cysteine at the C-terminus for labeling purposes.

KID-AB: TDSQKRREILSRRPSYRKILNDLSSDAPGVC
KID-B: RRPSYRKILNDLSSDAPGVC
Peptide C: RRLSFFYILLDLYLDAPGVC

**Peptide Synthesis, Purification and Modification**

PPKID and control peptides used to determine affinity for GST-KIX, HisKIX, KIX variants and in competition experiments and structure determination were synthesized on a 25 µmol scale using Fmoc chemistry at the Keck Foundation Biotechnology Resource Laboratory, Yale University School of Medicine (New Haven, CT). Phosphoserine residues were introduced using an Fmoc-protected O-benzylphosphoserine derivative with standard coupling conditions. All peptides contained an N-terminal amine and C-terminal amide.

PPKID4° variants were synthesized on a 25 µmol scale on rink amide resin using Fmoc chemistry on a Symphony Peptide Synthesizer (Protein Technologies, Inc.) as indicated above. The peptides were cleaved on the synthesizer using a cocktail containing
95% trifluoroacetic acid, 1% triisopropylsilane, 1% phenol, 1% water, 1% thioanisol and 1% ethanedithiol. Peptides were then precipitated in cold diethyl ether.

Crude peptides were purified by reverse-phase HPLC on a Vydac semi-preparative C18 column (300 Å, 5 µm, 10 mm x 150 mm). Matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometry was used to confirm peptide identity before further modification. Fluorescein-conjugated derivatives were generated by reaction of purified peptides containing single C-terminal cysteine residues with a 10-fold molar excess of 5-iodoacetamidofluorescein in a 3:2 mixture of dimethylformamide:phosphate-buffered saline (DMF:PBS). Labeling reactions were incubated with rotation for 3-16 hours at room temperature. Fluorescein-labeled peptides were purified by reverse-phase HPLC as described, and characterized by MALDI-TOF mass spectrometry and amino acid analysis.

For circular dichroism studies, peptides were alkylated on the C-terminal cysteine to prevent aggregation. Purified peptides were incubated with a 10-fold molar excess of 2-bromoacetamide in a 1:1 mixture of (DMF:PBS). Labeling reactions were incubated with rotation for 2-6 hours at room temperature. Alkylated peptides were purified by reverse-phase HPLC as described, and characterized by MALDI-TOF mass spectrometry and amino acid analysis.

**Fluorescence Polarization**

*Single sample format*

Fluorescence polarization experiments were performed with a Photon Technology International QuantaMaster C-60 spectrofluorimeter at 25 °C in a 1 cm pathlength
Hellma cuvette. Serial dilutions of HisKIX were made in KIX storage buffer. An aliquot of fluorescently labeled peptide was added to serial dilutions to a final concentration of 25 nM and the sample was incubated for 30 min at 25 °C to allow the binding reaction to reach equilibrium. Polarization was measured by excitation with vertically polarized light at a wavelength of 492 nm (10 nm slit width) and subsequent measurement of the fluorescence emission at a wavelength of 515 nm (10 nm slit width) for 30 s in the vertical and horizontal directions.

**Multi-well format**

Fluorescence polarization experiments were performed using an LJL Biosystems Analyst™ AD at 25 °C in Corning Costar non-treated 384 well flat-bottom black plates. Excitation and emission filters were specific for fluorescein; excitation filter-485 nm, emission filter-530 nm. Total volume per well was 32.5 µl and measurements were taken at a Z-height of 1.5 mm. Serial dilutions of KIX were made in KIX storage buffer. Fluorescently labeled peptide was added to the serial dilutions to a final concentration of 25 nM and the binding reaction was incubated for 30 min at 25 °C.

The polarization data were fit using Kaleidagraph v3.51 software to equilibrium binding equation (1.1), derived from first principles.

\[
(1.1) \quad P_{\text{obs}} = P_{\text{min}} + \frac{(P_{\text{max}} - P_{\text{min}})/(2[\text{peptide}^{\text{Flu}}])}{([\text{peptide}^{\text{Flu}}] + [\text{target protein}] + K_d - (((([\text{peptide}^{\text{Flu}}] + [\text{target protein}] + K_d)^2 - 4[\text{peptide}^{\text{Flu}}][\text{target protein}])^{0.5})}
\]

In this equation, \( P_{\text{obs}} \) is the observed polarization at any KIX concentration, \( P_{\text{max}} \) is the maximum polarization value, \( P_{\text{min}} \) is the minimum observed polarization value, and \( K_d \) is
the equilibrium dissociation constant. Measurements from three independent sets of samples were averaged for each dissociation constant determination. For plots of fraction of fluorescently labeled peptide (peptide$^{\text{Flu}}$) bound as a function of target protein concentration, polarization values were converted to fraction of peptide$^{\text{Flu}}$ bound using the $P_{\min}$ and $P_{\max}$ values derived from equation (1.1), and the fraction of peptide$^{\text{Flu}}$ bound data were fit to equilibrium binding equation (1.2) using Kaleidagraph v3.51 software.

$$\theta_{\text{obs}} = \frac{(1/(2[\text{peptide}^{\text{Flu}}]))([\text{peptide}^{\text{Flu}}] + [\text{target protein}] + K_d - (([\text{peptide}^{\text{Flu}}] + [\text{target protein}] + K_d)^2 - 4[\text{peptide}^{\text{Flu}}][\text{target protein}])^{0.5}}{[\text{target protein}] + K_d}$$

In equation (1.2), $\theta_{\text{obs}}$ is the observed fraction of peptide$^{\text{Flu}}$ bound at any KIX concentration and $K_d$ is the equilibrium dissociation constant.

**Fluorescence Polarization Competition**

For competition experiments, serial dilutions of KID-AB$^P$ were incubated with 1.5 $\mu$M or 3 $\mu$M HisKIX and 25 nM fluorescein-labeled PPKID4$^P$ or PPKID6$^U$ (peptide$^{\text{Flu}}$) for 60 min at 25 °C, respectively. Observed polarization values were converted to fraction of peptide$^{\text{Flu}}$ bound using experimentally determined $P_{\min}$ and $P_{\max}$ values corresponding to the polarization of samples containing 25 nM peptide$^{\text{Flu}}$ alone and peptide$^{\text{Flu}}$ with 1.5 $\mu$M or 3.0 $\mu$M HisKIX, respectively. The fraction of peptide$^{\text{Flu}}$ bound data were fit to equation (1.3) using Kaleidagraph v3.51 software to determine the IC$_{50}$ value.

$$\theta_{\text{obs}} = \frac{(\theta_{\max} - \theta_{\min})/(1 + ([\text{competitor}]/\text{IC}_{50})^{\text{slope}}))}{\theta_{\min}}$$
In equation (1.3), $\theta_{obs}$ is the observed fraction of peptide$^{\text{Flu}}$ bound at any KID-AB$^p$ concentration, slope is defined as the slope at the inflection point and IC$_{50}$ is the concentration of KID-AB$^p$ that reduces binding of peptide$^{\text{Flu}}$ by 50%.

**Circular Dichroism**

Experiments were performed on an Aviv Spectrometer Model 202 using a 1 cm pathlength Hellma cuvette. Samples contained peptides at 10 $\mu$M in a buffer containing 10 mM potassium phosphate (pH 7). Wavelength scans were performed from 260-190 nm with a 1 nm step size and an averaging time of 2 sec. The data represents an average of three scans per sample. Observed ellipticity was converted to mean residue ellipticity using equation (1.4).

(1.4) $\text{MRE} = \frac{\theta}{(10^4 c n l)}$

Where $\theta$ is the observed ellipticity, $c$ is the concentration in moles/liter, $n$ is the number of amino acids in the peptide and $l$ is the pathlength of the cuvette in centimeters. The MRE at 222 nm was used to calculate the fraction of the peptide that is $\alpha$-helical (f) using equation (1.5).

(1.5) $f = \frac{\text{MRE}_{222} - 500}{[(4 \times 10^4)^*(n-4/n)] - 500}$
References


(64) Rutledge, S. E. In *Chemistry Department*; Yale University: New Haven, 2003.


Chapter 2  Transcriptional Activation Potential of PPKID Ligands
Introduction

Regulation of transcription

It is estimated that the human genome is made up of approximately 30,000 genes\(^1\) corresponding to upwards of 66,000 gene products, by some estimations, due to alternative splicing.\(^2\) While “housekeeping” genes are always expressed, many genes are activated only in response to specific stimuli and cellular conditions. Nature has evolved a system to regulate transcription of genes through a network of signaling pathways leading to the activation or repression of a specific gene. In a simplified model, a transcription factor binds to the promoter of a gene to “turn it on”. Likewise, repressors can bind the promoter to “turn off” the gene. Despite the number of checks and controls used to ensure normal regulation of gene expression, the system can go awry leading to a variety of human diseases.

Aberrant transcriptional activities due to mutation of the transcription factor itself or regulators of the transcription factor are associated with a number of human diseases such as diabetes and cancer. For example, mutations in hepatocyte nuclear factors HNF-1\(\alpha\) and HNF-4\(\alpha\), transcription factors involved in cellular differentiation and metabolism, are known to impair insulin secretion and cause type 2 diabetes.\(^3\) HNF-4\(\alpha\) is an upstream regulator of HNF-1\(\alpha\), which activates the expression of liver-specific genes, including those involved in glucose metabolism, and potentially other genes in the kidney, intestine and pancreatic islets. In this case, a mutation in either the transcription factor or its regulator has the same result, illustrating the complex role of transcription factors in human disease.
Cancer is a second and more well-known example of atypical transcriptional activity leading to altered gene expression. In most cases, altered gene expression results from mutations in the signaling pathways that either activate or repress transcription factors. Significant interest exists in exploiting this fact to develop transcription therapy for cancer by antagonizing oncogenic activities. Two methods of therapeutic intervention have been suggested: first, directly altering the activity of the transcription factor and second, modulating critical target genes of the specific transcription factor. Towards this end, there is great interest in designing artificial transcription activators as therapeutics.

**Mechanism of activation**

Eukaryotic transcriptional activators stimulate gene expression primarily by recruitment of the general transcription machinery to the promoters of the specific genes they regulate. These transcription factors are modular in nature, containing a DNA binding domain for localization to the promoter, and an activation domain for recruitment of the transcriptional machinery. The functions of these two domains are separable; domains can be swapped among different activator proteins or replaced altogether with non-natural counterparts to obtain molecules with novel activation activities.

This model of activation by recruitment is supported by activator bypass experiments. In these experiments the need for an activating region is removed by tethering a DNA-binding domain to a component of the transcriptional machinery. As a result, high levels of transcription are often achieved with these constructs. Alternatively, the interaction between the transcriptional machinery and the activation domain can be replaced with an unrelated protein-protein interaction. With one protein partner tethered
to the DNA and the other protein tethered to a component of the transcriptional machinery, transcription function is retained. Activator bypass experiments, along with other observations that DNA binding is required for activation and the transcription machinery is indeed recruited to the promotor upon activation, verify the role of the activator as a recruiter of the transcription machinery.\(^5\)

The modularity and exchangeable functions of transcription factors is demonstrated by domain swap experiments. For example, the activation domain of yeast Gal4 attached to the DNA-binding domain of bacterial repressor LexA can activate a gene in yeast bearing LexA binding sites.\(^7\) In fact, a surprising number of DNA-binding and activating domains from bacteria, viruses and eukaryotes can function in many different organisms. Because they retain their function in many organisms, the Gal4 DNA-binding domain and the viral VP16 activation domain are often used to test the potential of artificial activation domains and DNA-binding domains, respectively, to activate transcription. Therefore, it is reasonable to believe that synthetic DNA-binding or activation domains will function in a variety of organisms similar to their natural components.

*Artificial transcription factors*

The development of fully artificial activators with both synthetic DNA-binding and activation domains is an important goal in chemical biology. While there has been considerable success in the development of novel DNA-binding molecules, the development of artificial activation domains lags behind.\(^8\) One explanation for this is that the DNA-binding domains of transcriptional regulators are well characterized from a
structural standpoint, whereas regulatory modules are less well-defined structurally and only a few have been characterized.

In the design of artificial transcription factors a modular approach is used, where the DNA-binding and activation components are developed individually. Fusing the artificial component to a heterologous partner - usually the yeast Gal4 DNA-binding domain or viral VP16 activation domain - allows one to test the activation potential of the new component and compare it to natural activators. Ultimately, these artificial activators will function to dissect genome-wide transcriptional pathways, monitor various cellular phenomenon as well as for therapeutic benefits.⁹

**Natural DNA-binding domains**

The DNA-binding domains of transcriptional regulators consist of a wide variety of structural folds and many domains are well characterized from both a structural and functional standpoint. As shown in Figure 2.1, examples of these domains include the homeodomain, bZIP, helix-loop-helix, zinc finger, zinc cluster and Ig-fold families of proteins. Many of these motifs contain an α-helix that inserts into the major groove of the DNA to recognize their specific binding site. As a result of this method of binding, these natural DNA-binding domains typically display high affinity and specificity for their target sequences.

**Artificial DNA-binding domains**

Initial focus in the artificial activator arena was placed on the DNA-binding domains because of the specificity with which they bind and regulate genes.⁶ The structures of natural DNA-binding domains (Figure 2.1) provide a starting point for
Figure 2.1 Natural DNA-binding domains. Structures of DNA-binding motifs (pink) bound to their cognate DNA sites (gray). Engrailed homeodomain (PDB 3hdd), Gcn4 bZIP (PDB 1ysa), Max helix-loop-helix (PDB 1an2), Zif268 zinc finger (PDB 1aay), Gal4 zinc cluster (PDB 1d66) and p53 Ig fold (PDB 1tup).
design of synthetic DNA-binding domains. However, the structural knowledge of natural DNA-binding domains does not provide a clear DNA recognition code which could be used to rationally design sequence-specific DNA-binding motifs.⁶

Advances in the selection and design of both peptidic and non-peptidic DNA-binding domains have overcome this lack of a simple recognition code for DNA. Barbas and colleagues have designed a selection strategy¹⁰ to identify zinc-finger based DNA-binding domains capable of targeting unique promoter sequences up to 18 base pairs in length. Furthermore, these zinc finger proteins are successful in controlling gene expression in cultured cells.¹⁰ In addition, there are a number of non-peptidic DNA-binding domains that function comparable to natural domains. For example, triplex-forming oligonucleotides (TFOs) bind to the major groove of double-stranded DNA through the formation of Hoogsteen or reverse Hoogsteen hydrogen bonds. TFOs have also been used in cell culture to inhibit and activate transcription initiation as well as inhibit transcription elongation.¹¹ Peptide nucleic acids (PNAs) are also triplex-forming molecules, however these DNA mimics contain a pseudopeptide backbone that displaces one strand of DNA forming a triplex structure with a 2:1 (PNA:DNA) ratio.¹² PNAs were shown to activate transcription in vitro and in cell culture by creating an artificial promoter.¹¹ More recently, Dervan and colleagues have designed polyamides which bind in a sequence specific manner to the minor groove of DNA¹³ (Figure 2.2). These polyamides contain combinations of pyrrole, hydroxypyrrole and imidazole pairs that can recognize all canonical DNA base pairs. The polyamide DNA-binding domains can be used alone to inhibit transcription or in combination with an activation domain to activate transcription.
**Figure 2.2** Non-natural DNA-binding domain. The polyamide DNA-binding domain (pink), PDB 1cvy, contains combinations of pyrrole, hydroxypyrrole and imidazole pairs that bind the minor groove and can recognize all canonical DNA base pairs.
**Natural activation domains**

While many DNA-binding domains of transcriptional regulators are well characterized from both a structural and functional standpoint, regulatory domains (both activators and repressors) contain little secondary structure and only a few have been characterized structurally. Instead of structural motifs, activation and repressor domains are characterized by the prevalence of certain amino acid residues. Activation domains contain acid-rich, glutamine-rich and/or proline-rich regions, while repressor domains contain alanine-rich or positively charged regions. Based on their composition of negatively charged and hydrophobic regions, it is suggested that activation domains are adhesive-like surfaces used to bind the components of the transcriptional machinery. Activation domains often lack distinct structure, however some activation domains, namely the acidic activators, adopt an amphipathic helix structure upon binding their target. For example, the most robust and universal activation domain, VP16 from the herpes simplex virus, contains multiple acidic and hydrophobic residues which form an amphipathic helix upon binding to its target hTAFII31.

Interestingly, not all activators appear to function via the same mechanism. Bentley and colleagues previously classified transcriptional activation domains into three distinct groups: Type I, Type IIA and Type IIB. Type I activation domains stimulate transcription initiation only; Sp1 (glutamine-rich) and CTF (proline-rich) fall into this category. Type IIA stimulate transcription elongation and this group has just one identified member, HIV Tat. The third class, Type IIB, is made up of acidic activators including VP16, p53 and E2F1 and stimulates both initiation and elongation. One hypothesis attributes these functional differences to contacts with different general
transcription factors. This study found a correlation between those activators that stimulate elongation and those that bind TFIIH. Another study shows that bulky hydrophobic residues within the activation domain are necessary for elongation, whereas acidic residues within the activation domain are primarily responsible for initiation.

*Artificial activation domains*

Natural activation domains lack a defined structure, unlike the highly-structured DNA-binding domains, therefore there is little structural information to guide the design of artificial activation domains. One of the first examples of an artificial activation domain was the peptide AH. This peptide was simply designed by choosing a sequence that could form a four-turn, negatively-charged, amphipathic helix. The amino acid sequence of AH, containing glutamine, glutamic acid and leucine residues, is arranged so that an amphipathic α-helix with one negatively charged face and one hydrophobic face is formed. In yeast, AH fused to the Gal4 DNA-binding domain activated transcription at a level 20% of that obtained with the natural Gal4 activation domain. More recent studies demonstrate that AH activates transcription when linked to a polyamide DNA-binding domain, creating one of the first transcription factors made up entirely of artificial components. In addition to rational design, a second method to generate peptidic activation domains uses genetic screens in conjunction with libraries of random peptide sequences to identify activation domains that interact with specific components of the transcriptional machinery.

The first non-peptidic activators discovered were RNA-based, but to date there are only two examples of RNA activators. The first was identified during a study using a yeast three-hybrid with an unconstrained library to find RNA ligands for the yeast Snpl
protein. The second activator was selected from a library of RNA stem-loop structures containing a variable 10 nucleotide loop attached to an invariable stem structure. The most active RNA aptamer (RNA #7) identified from the library activates transcription more than 400-fold over the control; this is approximately 15-fold less than the strong yeast Gal4 activator. Interestingly, RNA #7 is a better activator than the first RNA activator, which was selected from an unconstrained library.

Mapp and colleagues recently designed the first small molecule transcriptional activation domain. This activation domain was selected from a series of molecules based on the heterocyclic isoxazolidine scaffold modified with functional groups commonly found in natural activation domains. The most transcriptionally potent member of the series isoxazolidine 4 contains a combination of polar and hydrophobic groups. This small molecule activates transcription at a level similar to the viral activator VP16, as determined by an in vitro transcription assay.

**Correlation between affinity and activation**

Many studies show there is a direct correlation between in vitro affinity of the activation domain for its target protein and the level of transcriptional activation. Interestingly, one study also found that this correlation holds when comparing the activation potential of a particular activation domain to its in vitro affinity for non-transcriptional proteins. In this study, the affinity of a series of activation domains including Gal4, Gcn4, VP16, p53, HIV Tat, and the artificial domain AH for a set of transcriptional proteins TBP, TAF40 and TFIIB and non-transcriptional proteins Sug1p and lysozyme was determined. In addition, the activation potential of each activation domain was determined. The activation domains have a wide range of activation
potentials that is mirrored by their affinity for both transcriptional and non-transcriptional proteins. Interestingly, the activation domains that have the highest activation potentials, also have increased affinity for all of the proteins tested, whereas domains with low activation potential have comparatively less affinity for the proteins tested.

There is also contrasting evidence indicating that transcriptional activation levels do not always correlate with binding affinity. Mapp and colleagues screened peptide libraries (8 residue library with four randomized positions) to identify ligands for Gal11. Gal11 is a component of the yeast mediator complex, which is thought to act as a bridge between DNA-bound activators and the RNA pol II. Studies show that Gal11 activates transcription when fused to a DNA-binding domain making it an attractive target for artificial transcription factors. Three classes of ligands for Gal11 were identified: amphipathic peptides (similar to endogenous activators), hydrophobic peptides and positively-charged peptides. Binding data suggest that these three classes of ligands target at least three distinct binding surfaces of Gal11. While the binding affinities of these ligands are in the low micromolar range, they have a varied ability to activate transcription.

**CBP as a transcriptional coactivator**

Coactivators are a set of factors that respond to activators and may precede the recruitment of the polymerase to the promoter. The term coactivator is used because these molecules are thought to serve as molecular bridges between the activator and the transcription machinery; CBP (CREB-binding protein) and its paralog, p300, are two such proteins. CBP and p300 also have demonstrated acetyl transferase activity, which may help activators in binding their cognate sites on DNA, in addition to assisting
directly in recruiting the transcription machinery. Also, there is evidence that CBP and p300 may acetylate histone tails, which may help the transcription machinery to access the promoter.9 For these reasons, CBP is a crucial protein in regulation of transcription, binding to a number of activators including: CREB, c-myb, HIV Tat, HTLV Tax, c-jun, fos and p53. In addition, CBP may interact with one or more parts of the transcriptional machinery. Some reports classify CBP as part of the mediator complex, a complex first identified in yeast and believed to be a part of the transcription machinery.5 The key role of CBP, and its paralog p300, in transcription regulation is underscored by the findings that mutations in the CBP gene are described in various types of cancer and in Rubinstein–Taybi syndrome, both of which are synonymous with aberrant transcriptional activity.25

Miniature protein ligands for the KIX domain of CBP

In Chapter 1, the protein grafting strategy was used to identify a set of molecules that recognize the shallow surface of the transcriptional coactivator CBP with high nanomolar to low micromolar affinity.26 These molecules were identified from a phage library of phosphorylated or unphosphorylated aPP variants containing the α-helical binding epitope found on the KID domain of transcription factor CREB. The highest affinity phosphorylated (PPKID4\(^P\)) and unphosphorylated (PPKID6\(^U\)) CBP KIX ligands were characterized in detail. Both PPKID4\(^P\) and PPKID6\(^U\) bind CBP KIX as well as the phosphorylated KID domain of CREB (KID-AB\(^P\)), with equilibrium dissociation constants of 515 ± 44 nM and 1.5 ± 0.1 μM, respectively. Moreover, both PPKID4\(^P\) and PPKID6\(^U\) compete, in fluorescence polarization experiments, with KID-AB\(^P\) for binding to CBP KIX and exhibit high specificity for CBP KIX over proteins that bind
hydrophobic or \( \alpha \)-helical molecules, such as carbonic anhydrase and calmodulin.\(^{26} \) A more detailed analysis of the CBP KIX residues involved in PPKID binding reveals that despite conformational differences the \( \alpha \)-helix of PPKID\(^4\) binds in the same hydrophobic pocket of CBP KIX and with a similar orientation as the B helix of KID-AB\(^p\). However, PPKID\(^6\)\(^U\) binds an overlapping yet distinct region of CBP KIX. The distance separating the CBP KIX residues critical for binding this ligand suggests that PPKID\(^6\)\(^U\) binds in a conformation, where the polyproline helix makes additional contacts with CBP KIX. Binding affinity data for CBP KIX and PPKID\(^4\)\(^p\) variants, where residues forming the putative hydrophobic core were substituted with alanine, support a model in which PPKID\(^4\)\(^p\) folds into an aPP-like conformation upon binding to CBP KIX. By virtue of their ability to bind CBP, a transcriptional coactivator protein that acts as a conduit between gene-specific activators and the basal transcription machinery, we hypothesized that PPKID\(^4\)\(^p\) and PPKID\(^6\)\(^U\) might function as artificial activation domains when fused to the Gal4 DNA-binding domain. These studies may also provide information to understand how differences in binding orientation of the activation domain affect transcription activation and provide key details for a general mechanism of transcription activation.

**Results**

**Experimental goals**

The primary objective of this study was to determine the transcriptional activation potential of selected PPKID ligands in cultured human cells. To investigate the activation potential of these ligands, we prepared a series of mammalian expression
plasmids in which the gene encoding the CBP KIX ligands KID-AB, PPKID4 or PPKID6 was fused to the C-terminus of the gene encoding the Gal4 DNA-binding domain (GAL4 DBD). The level and pathway of transcriptional activation of PPKID4 and PPKID6 were compared to that of the natural CBP KIX ligand (KID-AB). Finally the potential of PPKID4\textsuperscript{5} to inhibit CBP-dependent transcription was determined.

*Constructs for transcription reporter assay*

Plasmids Gal4-KIDAB, Gal4-PPKID4 and Gal4-PPKID6 were created by inserting the genes encoding CBP KIX ligands KID-AB, PPKID4 and PPKID6 as C-terminal fusions to the gene encoding the Gal4 DBD through molecular cloning techniques (Figure 2.3). These expression plasmids are derivatives of the previously reported expression vector pAL\textsubscript{1,23} which directs the expression of the Gal4 DBD (residues 1-147). Using these constructs with an appropriate reporter system, the potential of PPKID peptides to activate transcription can be measured in a human cell line. In addition to the Gal4-PPKID and Gal4KIDAB constructs, several other plasmids were used in the transcription reporter assay including: G5B luciferase reporter, pRL and pBluescript. The G5B luciferase reporter vector\textsuperscript{23} (a gift from John Frangioni, Harvard Medical School) contains five Gal4 binding sites upstream of a firefly luciferase gene. As an internal control pRL-null, a vector containing a promoterless *Renilla* luciferase gene, corrects for transfection efficiency. The pBluescript vector acts as carrier DNA to protect other DNA from degradation during transfection. These three vectors were co-transfected with one of the Gal4-peptide constructs to determine the transcriptional activation potential of that particular peptide. As shown in Figure 2.4, the Gal4-peptide fusions bind to Gal4 sites on the reporter. If the peptide can recruit endogenous CBP and
**Figure 2.3** Plasmid map of the pAL₁ expression vector (adapted from Voss et al.). The pAL₁ vector encodes residues 1-147 of the Gal4 DNA-binding domain. The DNA and amino acid sequences of the multiple cloning site are shown. Peptides expressed as C-terminal fusions to the Gal4 DNA-binding domain are inserted between the SalI and BamHI restriction sites.
Gal4 DNA Binding Domain a.a. 1-147

pAL1

Multiple Cloning Site

Gal4 DNA Binding Domain a.a. 1-147

Bgl II

CMV IEP

KanR

5' TCG CCG GGA TTC GGA TCA GGG TCG ACG GGT ACC GGA TCC TGA 3'

S 147 P E F G S G S T G T G S *

Gal4 DBD

Sal I Kpn I BamHI Stop

**Figure 2.4** Schematic of transcription reporter assay. PPKID ligands are fused to the Gal4 DNA-binding domain, which binds to Gal4 binding sites in the promoter. Ligands capable of recruiting CBP, which in turn can recruit the transcription machinery, will result in transcription of the luciferase gene. The activity of the luciferase gene product in the cell lysate can be detected with beetle luciferin, which is converted to oxyluciferin and light in the presence of firefly luciferase.
5X Gal4 binding sites

Gal4 DBD - PPKID

CBP

Basal Transcription Machinery

mRNA

Luciferase Gene

+ Light

Luciferase

5X Gal4 binding sites

mRNA
the rest of the transcription machinery, then the luciferase gene will be transcribed. Therefore, the ability of PPKID4\(^p\) and PPKID6\(^u\) to mimic the role of KID-AB\(^p\) and recruit the transcriptional machinery will be measured by the amount of luciferase in the cell lysate.

*Expression of constructs in HEK293 cells*

The human embryonic kidney cell line 293 (HEK293) was used for all cell culture experiments. HEK293 is an adherent epithelial cell line that is commonly used for transcription experiments involving CREB.\(^{23,27}\) The expression vectors contain a cytomegalovirus (CMV) promoter, which is a strong, constitutive promoter. Constructs were transfected into HEK293 cells using Superfect reagent (Qiagen). The Superfect reagent is an activated dendrimer that assembles DNA into compact structures; these compact structures bind to the cell surface and are taken up into the cell by non-specific endocytosis. Expression of the Gal4-peptide constructs in HEK293 cells were verified by western blotting (Figure 2.5) using an antibody for the Gal4 DBD. The Gal4 DBD alone has an apparent molecular weight of 18 kDa\(^{33}\) while Gal4-KIDAB, Gal4-PPKID4 and Gal4-PPKID6 constructs have an apparent molecular weight of 22 kDa.

*Phosphorylation of constructs in HEK293 cells*

As reported in Chapter 1, PPKID4 only binds with high affinity to CBP (*in vitro*) when phosphorylated. In these experiments, this post-translational modification must be initiated in cell culture. By design, all PPKID peptides contain the recognition sequence for protein kinase A (PKA). PKA is endogenous to mammalian cells, however it requires a stimulus in order to be active. PKA is activated when its regulatory subunit dissociates
Figure 2.5  Expression of Gal4 DNA-binding domain constructs. Western blot of Gal4 DNA-binding domain fusion to PPKID and KID-AB constructs in HEK293 cells. Apparent molecular weight of Gal4 DNA-binding domain is 18 kDa, PPKID and KID-AB fusions are 22 kDa.
20 kDa

none  Gal4  PP  PP  KID
DBD  KID6  KID4  AB
from the catalytic subunit; the free catalytic subunit is the active form of the kinase. Therefore, the concentration of activated PKA enzyme needs to be increased in the cells to effectively phosphorylate the PPKID peptides.

Here, two different methods were used to phosphorylate the PPKID peptides expressed in cells. The first phosphorylation method transfects a plasmid encoding the gene for the catalytic subunit of PKA, therefore the constitutively expressed PKA would not require activation. A series of experiments were performed where 5 – 50 ng of the PKA plasmid were co-transfected with 5 ng of the Gal4-peptide constructs. As the amount of PKA transfected increased, so did the level of transcription activation by Gal4-peptide constructs. Unfortunately, the level of activation was not consistent from experiment to experiment. This inconsistency may result from competition among all of the expression plasmids. A number of plasmids were transfected together, many of them driven by the strong CMV promoter, causing all to compete for the same cell resources for transcription and expression. This may result in varying amounts of PKA expressed in the cells and therefore varying levels of phosphorylation resulting in varying levels of active PPKID4.

The second method is to stimulate the cAMP signaling pathway, leading to the activation of endogenous PKA. The small molecule forskolin activates the cAMP pathway increasing the cellular concentration of cAMP. cAMP then binds the regulatory subunit of PKA and releases the active form of PKA.28 Cells were stimulated 30 hours post-transfection by adding forskolin to the media. Stimulation of the cAMP pathway starts within 30 minutes of forskolin treatment and levels off after 4 hours.28 The cells were incubated with forskolin for 6 hours and then harvested. Various concentrations of
forskolin from 5 - 100 µM were tested for their effect on transcriptional activation. Concentrations above 20 µM forskolin resulted in a decrease of luciferase activity and at 100 µM luciferase activity was essentially zero indicating that high levels of forskolin are detrimental to the cell. Transcriptional activation was highest with 5 µM forskolin and resulted in highly reproducible transcriptional activation. Therefore this concentration was used in subsequent studies to phosphorylate PPKID and KID-AB constructs.

Phosphorylation of PPKID4 and KID-AB (the positive control) through stimulation with forskolin leads to an observed increase in transcriptional activation. Unfortunately phosphorylation of the peptides could not be verified by western blot. Many commercially available antibodies recognize a phosphoserine in the context of a specific protein, including phosphoserine 133 of CREB. However, our PPKID molecules do not present the PKA recognition site in the same context as in CREB because the surrounding residues come from the aPP scaffold. A western blot was performed on HEK293 cell lysate from cells transfected with Gal4-PPKID or Gal4-KIDAB constructs using an antibody (Sigma P3430) that recognizes phosphoserine in any context, except conformationally hindered phosphoserine residues. Cells were stimulated with forskolin for 2 hours before harvesting, to promote cAMP-dependent phosphorylation and minimize dephosphorylation by phosphatases. Multiple bands were present by western blotting, indicating that phosphoserine was detected in various proteins expressed in the cells. To detect phosphoserine in the PPKID and KID-AB constructs, a western blot was performed on proteins that were immunoprecipitated with an antibody for the Gal4 DBD. While the PPKID and KID-AB constructs were clearly pulled down from the lysate, phosphoserine was not detected. However, it is possible that this particular antibody does
not recognize the phosphoserine in the context it is presented in these molecules. In order to verify the presence of phosphoserine a more inclusive search of many phosphoserine antibodies would have to be undertaken.

*Transcriptional activation by PPKID*

We investigated the ability of KID-AB, PPKID4, and PPKID6 to activate transcription under conditions where no external stimulation of phosphorylation of the peptides took place (Figure 2.6). In the absence of phosphorylation, PPKID4 and KID-AB possess only low affinity for CBP KIX *in vitro* and thus would not be expected to recruit significant amounts of CBP to the Gal4 promoter in cells. As expected under these conditions, PPKID4 and KID-AB did not activate transcription to any appreciable extent compared to transcription activation of the control plasmid pAL1 (this plasmid expresses the Gal4 DBD only). Unexpectedly, however, PPKID6 also failed to activate transcription under these conditions despite possessing low micromolar affinity for CBP KIX in the absence of phosphorylation. The overall levels of firefly luciferase activity in cells transfected with Gal4-KIDAB, Gal4-PPKID4, and Gal4-PPKID6 were typically between 50 and 90% of the levels obtained in cells transfected with pAL1.

Next, we investigated the ability of KID-AB, PPKID4 and PPKID6 to activate transcription under conditions where phosphorylation of these ligands is stimulated. As phosphorylation of KID-AB and PPKID4 dramatically increases their *in vitro* CBP KIX-binding affinity, we expected that phosphorylation of these ligands will also increase their ability to recruit CBP and thus their transcriptional potency in HEK293 cells. When experiments were conducted in the presence of forskolin, KID-AB and PPKID4 activated transcription 7.5-fold over basal levels. Phosphorylation of PPKID6 also increases its *in
**Figure 2.6** Transcriptional activation by PPKID peptides and KID-AB in HEK293 cells. For each peptide, an expression vector was constructed which codes for the peptide fused to the c-terminus of the Gal4 DNA-binding domain (residues 1-147). HEK293 cells were transfected with 5 ng of each construct and assayed for activation. Where phosphorylation is indicated, 5 µM forskolin was added to culture media 6 hours before harvesting cells. Firefly luciferase values were normalized to an internal control (luciferase values from promoterless Renilla luciferase vector) to correct for transfection efficiency. Fold activation represents normalized luciferase values relative to values for the Gal4 DNA-binding domain alone under the same conditions. Bars and standard error represent the results from at least 3 independent trials.
fold increase in transcription

<table>
<thead>
<tr>
<th>Protein</th>
<th>Fold Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal4 DBD</td>
<td>1</td>
</tr>
<tr>
<td>KID6</td>
<td>1</td>
</tr>
<tr>
<td>KID6P</td>
<td>2</td>
</tr>
<tr>
<td>PP</td>
<td>1</td>
</tr>
<tr>
<td>PP</td>
<td>1</td>
</tr>
<tr>
<td>KID4</td>
<td>5</td>
</tr>
<tr>
<td>KID4P</td>
<td>8</td>
</tr>
<tr>
<td>AB</td>
<td>7</td>
</tr>
<tr>
<td>ABP</td>
<td>8</td>
</tr>
</tbody>
</table>
vitro CBP KIX-binding affinity. However, with only 2-fold higher *in vitro* affinity of the phosphorylated over the unphosphorylated peptide for CBP KIX, it was not clear whether PPKID6\(^\text{p}\) would be capable of activating transcription when PPKID6\(^\text{u}\) could not. However, in the presence of forskolin, PPKID6 activated transcription 2.5-fold over the pAL\(_1\) control. Increasing the amount of Gal4-KIDAB, Gal4-PPKID6 and Gal4-PPKID4 plasmids transfected (10 - 20 ng) did not increase the level of transcription. In fact, with increasing amount of plasmid transfected the level of transcription decreased slightly.

*Activation pathway of PPKID*

Transcriptional activation by CREB KID occurs via recruitment of CBP to promoters where CREB is bound.\(^{29}\) However in cells, transcription can be activated by a variety of different pathways. Therefore, it was of great interest to investigate whether the observed transcription activation potential of PPKID4\(^\text{p}\) and PPKID6\(^\text{p}\) is dependent on recruitment of CBP. Towards this end, we compared the transcription potential of PPKID4\(^\text{p}\) and PPKID6\(^\text{p}\) to KID-AB\(^\text{p}\) in the presence and absence of the E1A oncoprotein, which is known to sequester cellular p300/CBP activity, and exogenous p300, a paralog of CBP.

*Effects of oncoprotein E1A*

First we compared the transcription potential of PPKID4\(^\text{p}\) and KID-AB\(^\text{p}\) in the presence and absence of E1A oncoprotein. The ability of CBP to stimulate transcription is inhibited by the adenoviral E1A oncoprotein through direct protein-protein interactions.\(^{30}\) Increasing amounts of the plasmid CMV/E1A\(^{23}\) were co-transfected with 5 ng of Gal4-KIDAB, Gal4-PPKID4 or Gal4-PPKID6. Compared to transcription levels in
the absence of E1A, transcription decreased with increasing amounts of CMV/E1A. The
decrease in activation for PPKID4\textsuperscript{p} was identical to that for KID-AB\textsuperscript{p} (Figure 2.7). When
25 ng of CMV/E1A was transfected, transcription of PPKID4\textsuperscript{p} and KID-AB\textsuperscript{p} was
reduced to 65\% of levels in the absence of E1A. When 50 ng of CMV/E1A was
transfected, transcription levels were 50\% and with 100 ng of CMV/E1A transcription
levels dropped to 40\%.

Interestingly, the decrease in transcription also affected the level of transcription
of the control pAL\textsubscript{1}. Normalization of the transcription levels of PPKID4\textsuperscript{p} and KID-AB\textsuperscript{p}
to the levels obtained with pAL\textsubscript{1} under the same conditions resulted in no significant
decrease in fold transcription when E1A was expressed. Due to this discrepancy, we are
unable to draw any conclusions from this set of experiments concerning the pathway of
transcription activation by PPKID molecules.

Effects of transcriptional coactivator p300

Next, we investigated the transcription potential of PPKID4 and PPKID6 in the
presence and absence of exogenous p300, a paralog of CBP. Previous work has
demonstrated that in the presence of exogenous CREB, components of the transcriptional
machinery such as CBP are limiting\textsuperscript{,\cite{31}} Hence an increase in the CBP/p300 concentration
leads to an increase in transcriptional activation via the p300/CBP pathway. Thus, we
expected that the levels of transcription activation elicited by PPKID4\textsuperscript{p} and PPKID6\textsuperscript{p} in
the presence of exogenous p300 would be significantly higher than the levels of
transcription observed with only endogenous CBP/p300.

As expected, co-expression of exogenous p300 along with Gal4-KIDAB or Gal4-
PPKID4 and treatment with forskolin led to a 20-fold increase in the level of transcription
**Figure 2.7** Effect of E1A protein on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells. For each peptide, an expression vector was constructed which codes for the peptide fused to the c-terminus of the Gal4 DNA-binding domain (residues 1-147). HEK293 cells were co-transfected with 0 - 100 ng of an expression vector encoding E1A under control of the CMV promoter, 5 ng Gal4 DBD-peptide constructs and assayed for activation. Cells were treated with 5 µM forskolin 6 hours before harvesting. Firefly luciferase values were normalized to an internal control (luciferase values from promoterless Renilla luciferase vector) to correct for transfection efficiency. Percent activation represents luciferase values relative to values for PPKID4$^P$ or KID-AB$^P$ in the absence of E1A. Bars and standard error represent the results from at least 3 independent trials.
Amount of E1A plasmid transfected (ng)

Percent Activation

PPKID4^P

KID-AB^P
compared to the level observed in cells transfected with pAL₁, a 2.7-fold increase relative to the level observed in the absence of exogenous CBP/p300 (Figure 2.8). In the absence of forskolin, co-expression of exogenous p300 with Gal4-KIDAB led to a 7.5-fold increase in the level of transcription, equal to the level of transcription for phosphorylated KID-AB without exogenous p300. However in the absence of forskolin, co-expression of exogenous p300 and Gal4-PPKID4 did not activate transcription. Likewise, PPKID6<sup>p</sup> activated transcription 15-fold over basal levels in the presence of additional p300, a 6-fold increase relative to PPKID6<sup>p</sup>-dependent transcription mediated by endogenous CBP/p300. Somewhat surprisingly, addition of p300 also increased transcription activation by PPKID6<sup>U</sup> to levels 4.5-fold over basal transcription levels, whereas PPKID6<sup>U</sup> failed to activate transcription in the presence of only endogenous CBP/p300. Although PPKID6<sup>U</sup> only demonstrates modest activation, it nevertheless activates transcription via the CBP/p300 pathway. These results are consistent with a model in which PPKID4<sup>p</sup>, PPKID6<sup>p</sup> and PPKID6<sup>U</sup>, like CREB KID, activate transcription by recruitment of CBP/p300 to the promoter via the KIX domain.

**Inhibition of activation by PPKID4<sup>p</sup>**

The results presented above support a model in which Gal4-DBD fusion proteins containing PPKID4 or PPKID6 activate CBP-dependent transcription in a phosphorylation-dependent manner. This model suggests that PPKID4 and PPKID6, when separated from the Gal4-DBD, should inhibit transcription through competition with other proteins that interact with CBP. Therefore, to provide additional evidence of transcriptional activation via a CBP-dependent pathway, we analyzed the levels of transcription in forskolin-stimulated HEK293 cells transfected with 5 ng of Gal4-
Figure 2.8 Effect of p300 protein on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells. For each peptide, an expression vector was constructed which codes for the peptide fused to the c-terminus of the Gal4 DNA-binding domain (residues 1-147). HEK293 cells were co-transfected with 25 ng of an expression vector encoding full-length p300 under control of the CMV promoter and 5 ng Gal4 DBD-peptide constructs and assayed for activation. Where phosphorylation is indicated, cells were treated with 5 µM forskolin 6 hours before harvesting. Firefly luciferase values were normalized to an internal control (luciferase values from promoterless Renilla luciferase vector) to correct for transfection efficiency. Fold activation represents normalized luciferase values relative to values for the Gal4 DNA-binding domain alone under the same conditions. Bars and standard error represent the results from at least 3 independent trials.
KIDAB, Gal4-PPKID4 or Gal4-PPKID6 in the presence of increasing amounts (0 – 50 ng) of the pPPKID4 plasmid, encoding PPKID4 without the Gal4 DBD. Based on previous results, we hypothesize that the PPKID4\(^p\) activation domain will inhibit transcription of the phosphorylated PPKID and KID-AB ligands.

As expected, transcriptional activation by PPKID4\(^p\) was significantly reduced when a 2:1 ratio of pPPKID4:Gal4-PPKID4 was transfected in HEK293 cells with forskolin stimulation (Figure 2.9). Co-transfection of a 5-fold excess of pPPKID4 brought activation down to levels observed with the control plasmid pAL\(_1\). Activation by KID-AB\(^p\) was inhibited when a 5:1 ratio of pPPKID4:Gal4-KIDAB was transfected and activation was completely inhibited when a 10-fold excess of pPPKID4 was transfected. Due to its low level of activation, PPKID6\(^p\) activation requires only a 2-fold excess of pPPKID4 to return to basal levels. These results also support a model in which both PPKID4\(^p\) and PPKID6\(^p\) activate transcription in a CBP-dependent manner.

**Co-immunoprecipitation and pull-down of CBP**

Various methods were used to verify that the PPKID ligands were bound to CBP *in vivo* including co-immunoprecipitation and pull-down assays. In these experiments, non-denaturing lysis conditions were used when harvesting the cells and precautions were taken to inhibit phosphatases. Co-immunoprecipitation with an agarose conjugate of either anti-Gal4 DBD or anti-CBP failed to verify this interaction even for the positive control KID-AB\(^p\) under the conditions used. The Gal4 DBD constructs and CBP are precipitated with their respective antibodies, but CBP and the Gal4 DBD constructs did not co-immunoprecipitate, respectively. A second method for verifying the interaction of Gal4 DBD constructs and CBP was attempted using a pull-down assay. Syntheticl...
Figure 2.9 Effect of PPKID4\(^p\) on transcriptional activation by PPKID peptides and KID-AB in HEK293 cells. For each peptide, an expression vector was constructed which codes for the peptide fused to the c-terminus of the Gal4 DNA-binding domain (residues 1-147). HEK293 cells were co-transfected with 5 ng Gal4 DBD-peptide constructs and 0 – 50 ng of PPKID4 vector and assayed for activation. 5 µM forskolin was added to culture media 6 hours before harvesting cells. Firefly luciferase values were normalized to an internal control (luciferase values from promoterless Renilla luciferase vector) to correct for transfection efficiency. Fold activation represents normalized luciferase values relative to values for the Gal4 DNA-binding domain alone under the same conditions. Bars and standard error represent the results from at least 3 independent trials.
Fold Transcription Activation

Amount of PPKID4 vector transfected (ng)

PPKID6^p  PPKID4^p  KID-AB^p

0  10  25  50  0  10  25  50  0  10  25  50
prepared PPKID or KID-AB ligands were covalently bound to an agarose bead and incubated with the cell lysate. Western blotting was not able to detect any significant levels of CBP pulled-down by the PPKID-or KID-AB-agarose conjugate.

There are a couple of explanations regarding why these experiments may have been unsuccessful. First of all, CBP is not expressed at high levels in HEK293 cells and it is possible that a small amount of CBP bound to the PPKID constructs is not detectable by western blot methods. Secondly, the PPKID peptides require phosphorylation in order to interact with CBP. It was not clear that the PPKID constructs remain phosphorylated during the experiments and the lack of a reliable phosphoserine antibody adds to this problem.

Discussion

PPKID as transcriptional activators

Our results indicate that PPKID4\textsuperscript{p} and PPKID6\textsuperscript{p} can function as transactivation domains when fused to a heterologous DNA-binding domain. In fact, the level of transcriptional activation elicited by PPKID4\textsuperscript{p} is comparable to that of phosphorylated CREB KID, the natural protein from which the PPKID ligands were designed. This is not surprising based on data from Chapter 1 indicating that PPKID4\textsuperscript{p} binds an identical site on CBP KIX as KID-AB\textsuperscript{p}, and with comparable affinity. In addition, PPKID4\textsuperscript{p} and KID-AB\textsuperscript{p} give rise to similar transactivation profiles in the presence of additional p300, suggesting that transcription activation by both ligands occurs via the p300/CBP pathway. Further evidence that phosphorylated PPKID4 and KID-AB interact with the same site on CBP was verified by competition with the PPKID4 activation domain in
HEK293 cells. Transcription activation by KID-AB$^p$ is completely inhibited when a 5- to 10-fold excess of the PPKID4 activation domain is co-transfected. In the unphosphorylated state, neither KID-AB nor PPKID4 activates transcription. KID-AB$^u$ is a poor ligand for CBP KIX with an equilibrium dissociation constant greater than 116 µM, whereas PPKID4$^u$ has moderate affinity for CBP KIX with an equilibrium dissociation constant of 12.1 µM. It is possible that PPKID4 binds to CBP KIX in a different orientation in the absence of phosphorylation and this may explain the poor activation potential of PPKID4$^u$. Although the exact binding footprint of PPKID4$^u$ on CBP KIX has not been investigated, we hypothesize (based on evidence for PPKID6$^u$) that the phosphoserine is responsible for directing phosphorylated ligands into the CREB-binding site on CBP KIX.

Despite possessing high affinity for CBP KIX, PPKID6$^u$ does not activate transcription, and addition of p300 leads to only a modest increase in PPKID6$^u$-mediated transcription. PPKID6 was selected as an unphosphorylated ligand and as shown in Chapter 1, it utilizes an alternate mode of binding to CBP KIX to maximize affinity in the absence of the enthalpic driving force provided by interactions involving the phosphoserine moiety. Because residues critical for PPKID6$^u$ binding are widely scattered on the surface of CBP KIX, we hypothesize that PPKID6$^u$ is using both the α-helix and polyproline helix in binding to CBP KIX. Thus, it is possible that PPKID6$^u$ binding may impede interactions between CBP KIX and other proteins essential for activating transcription via the p300/CBP pathway. However, PPKID6$^p$ activates transcription via the p300/CBP pathway at a level 1/3 that of PPKID4$^p$ and KID-AB$^p$. In this case, phosphorylation may provide sufficient enthalpic driving force to direct
PPKID6\textsuperscript{p} into the CREB-binding site on CBP despite the fact that the unphosphorylated ligand was optimized to bind CBP KIX in an alternate conformation. Experiments with the PPKID4 activation domain further support a model where PPKID6\textsuperscript{p} competes for a similar site on CBP; a 2-fold excess of the PPKID4\textsuperscript{p} activation domain completely inhibits transcriptional activation by PPKID6\textsuperscript{p}.

\textit{Implications for unphosphorylated PPKID}

Recently the solution structure of CBP bound to the transactivation domain of c-Myb was solved.\textsuperscript{32} c-Myb is a constitutive activation domain that interacts with the same hydrophobic pocket of CBP KIX as CREB KID. Comparing the structures of c-Myb versus CREB KID bound to CBP KIX leads to interesting conclusions regarding structural and thermodynamic requirements for both constitutive and inducible activators. The structures reveal that c-Myb residues complement the hydrophobic surface of CBP KIX better than CREB KID and that the interactions made by c-Myb are more extensive. For example, both c-Myb and CREB KID have a leucine residue that is buried in a hydrophobic pocket on CBP KIX. However, a bend in the c-Myb helix, enables its leucine to further penetrate CBP KIX allowing it to make an additional contact with a leucine residue on CBP KIX located at the bottom of the hydrophobic pocket. These optimized interactions enable c-Myb to bind to CBP KIX at a level greater than unphosphorylated KID, but less than phosphorylated KID, allowing transcription to occur at a constitutive, albeit low, level. In light of these observations for c-Myb and CREB KID, it follows that the best way to select for an unphosphorylated ligand is to use a library that can optimize binding of the $\alpha$-helix to the hydrophobic pocket of CBP KIX. The PPKID Library 1 used for both phosphorylated and unphosphorylated selections
contains an invariable α-helix, therefore a new PPKID library with randomized residues on the CBP-binding portion of the α-helix could result in a high affinity unphosphorylated ligand for CBP and a more potent transcription activator.

**Comparison to other KIX ligands**

Previous studies on other examples of synthetic CBP-KIX-binding molecules suggested a correlation between CBP KIX binding affinity and transcriptional activation potential\(^{23}\). These studies identified a peptide, KBP 2.20, which bound CBP KIX with a \(K_d\) of 16 µM and activated transcription 40-fold via a p300/CBP-dependent pathway when fused to a Gal4 DNA-binding domain. Although our selected PPKID ligands have higher affinity for CBP KIX than KBP 2.20, the level of transcriptional activation they elicit is lower. Our results suggest that although CBP KIX binding affinity plays a role in determining the level of transcriptional activation mediated by a particular ligand, it is not the sole criterion. In fact, many factors may affect transcriptional activation potential including the lifetime of the ligand•CBP complex, the orientation of ligand binding, the resulting accessibility of CBP for binding to other essential transcription proteins, the rate of modification (phosphorylation), and susceptibility of the activator to degradation pathways in the cell. Our set of ligands could potentially be used to differentiate among these factors and provide key details for a general mechanism of transcription activation.

**Conclusions**

We have shown that PPKID\(^4\) and PPKID\(^6\) function as transcriptional activators much like the transcription factor CREB after which they were modeled. These ligands act as artificial activation domains, and have the potential to serve as tools in
understanding the mechanism of transcription activation. To expand the functionality of these molecules, a second-generation phosphorylated library based on PPKID4 is currently under construction. The goal of this new library is to improve the packing of the hydrophobic core of the scaffold resulting in a molecule that has high affinity for CBP KIX and is folded in the absence of CBP. Using recent data regarding PPKID6<sup>U</sup>, a new unphosphorylated library can be designed to optimize binding as well as improve packing of the hydrophobic core.

**Experimental**

**General**

**Materials**

Restriction enzymes, T4 DNA ligase, Vent DNA polymerase and associated reaction buffers were purchased from New England Biolabs (Beverly, MA) QIAfilter DNA Maxiprep and Miniprep kits, QIAquick gel extraction and PCR purification kits and QIAshredder columns were purchased from Qiagen (Valencia, CA). PhastGel media, Hybond-P PVDF membrane and ECF substrate was purchased from Amersham Biosciences (Piscataway, NJ). XL1 Blue electroporation competent *E. coli* cells were purchased from Stratagene (La Jolla, CA). Anti-Gal4 DBD (sc-510), anti-CBP (sc-7300) and RIPA buffer were purchased from Santa Cruz biotechnology (Santa Cruz, CA). Dulbecco’s modified Eagle medium (DMEM), fetal bovine serum (FBS), Trypsin-EDTA, L-glutamine, forskolin and anti-mouse IgG alkaline phosphatase conjugate (A-3562) were purchased from Sigma (St. Louis, MO). HEK293 cells were purchased from
American Type Culture Collection (Manassas, VA). The dual luciferase assay kit and promoterless *Renilla* luciferase plasmid was purchased from Promega (Madison, WI). 5X Gal4 firefly luciferase reporter plasmid and pAL1 vectors were gifts from John Frangioni (Harvard Medical School, Boston, MA). pBluescript SK+ vector was purchased from Stratagene (La Jolla, CA). Oligonucleotide synthesis was performed at the W. M. Keck Foundation Biotechnology Resource Laboratory, Yale University School of Medicine (New Haven, CT).

*Media and Buffers*

**LB media** contains 10 g tryptone, 5 g yeast extract and 10 g NaCl and 1 ml 1 M NaOH per 1 L. **LBKan agar** contains all components for LB media in addition to 15 g agar and 50 mg kanamycin per 1 L. **TE buffer** contains 10 mM Tris pH 8 and 1 mM EDTA. 1 L of DMEM complete contains 880 mL of DMEM, 100 mL fetal bovine serum and 20 mL 200 mM L-glutamine. Phosphate-buffered saline (PBS) contains 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄ and 0.24 g KH₂PO₄ per 1 L and is adjusted to pH 7.4. **TBS buffer** contains 20 mM Tris, pH 8 and 150 mM NaCl. **TBST buffer** is TBS buffer with the addition of 0.05% Tween-20. **1X SDS-loading buffer** contains 50 mM Tris•Cl pH 6.8, 2% sodium dodecyl sulfate (SDS), 0.1% bromophenol blue, 10% glycerol, 100 mM dithiothreitol (DTT) and protease inhibitors (1 Mini protease inhibitor tablet/10 mL (Roche)). **Towbin transfer buffer** contains 25 mM Tris, 192 mM glycine, 0.1% SDS, and 15% methanol. **Mini-gel running buffer** contains 3.03 g Tris, 15.5 g glycine, and 1 g SDS per 1 L.
Cloning of Gal4 DBD-peptide constructs

Peptides were cloned into the pAL1 vector\textsuperscript{33}, which encodes the Gal4 DNA-binding domain (residues 1-147), resulting in fusion of the peptide to the C-terminus to the Gal4 DBD, see Figure 2.3. Each PPKID peptide sequence was amplified from the pCANTAB_PPKID vector, resulting from phage display selections, by PCR using primers KID4-5 and KID4-3 for PPKID4 and KID6-5 and KID6-3 for PPKID6. SalI and BamHI sites were added to clone the insert into the pAL1 vector. KID4-5: 5’TCAGGGTGTACCCAGGACC3’, KID4-3: 5’TCAGGATCCC GGTGCAGTACAG3’, KID6-5: 5’TCAGGGTGTACCCAGGACC3’, KID6-3: 5’TCAGGATCCCAGGACC3’.

The KID domain of CREB (rat codons 121-151) was cloned as a C-terminal fusion to the Gal4 DBD. The KID domain was amplified from the pCANTAB_CREB vector (a gift from Stacey Rutledge) by PCR using primers CREB5 and CREB3. SalI and BamHI sites were added to clone the insert into the pAL1 vector. CREB5: 5’TCAGGGTGTACCCAGGACC3’, CREB3: 5’TCAGGATCCCAGGACC3’.

Primers were desalted on a Nap-10 column and concentrations were determined by absorbance at 260 nm. PCR samples contained 1 µl template, 1 µM each of primer pair, 300 µM dNTP’s in 1X ThermoPol buffer. Water was added to bring the final volume to 100µl. Immediately before the reaction was started, 1 µl Vent DNA polymerase was added to each sample. The inserts were amplified using the following PCR program: 1) 94 °C for 2 minutes, 2) 94 °C for 30 seconds, 3) 71 °C for 30 seconds, 4) 72 °C for 1 minute, 5) repeat steps 2-4 19x, 6) 72 °C for 3 minutes, 7) cool samples to
4 °C. The correct product was purified from contaminating products by separation on a 2% agarose gel and extracted from the agarose gel using a QIAquick Gel Extraction Kit. Purified insert was digested with 40 units each of BamHI and SalI in a total volume of 40 µl for 20 h at 37 °C. Digested DNA was purified using the QIAquick PCR purification kit.

The pAL₁ vector (10 µg) was digested with 40 units each of BamHI and SalI in a total volume 40 µL for 4 h at 37 °C. The digested vector was purified using the QIAquick PCR purification kit. Digested insert and vector were ligated using T4 DNA ligase at room temperature for 20 minutes. The ligation reaction (1 µl) was transformed into XL1Blue electroporation competent E. coli cells and plated on LBKan agar. Plasmid DNA was prepared from twenty clones and digested with KpnI to differentiate between the original vector and successful ligation products. The KpnI restriction site is destroyed when the insert sequence is correctly inserted into the pAL₁ vector. Clones resistant to KpnI digestion were sent for DNA sequencing. DNA from a correct CREB KID clone was prepared using a QIAfilter maxiprep kit to produce DNA for the reporter transcription assay. Resulting DNA was resuspended in TE buffer.

*pPPKID4 construct*

The Gal4 DBD portion of the pAL₁ vector was removed using the BglII and BamHI restriction sites. This also removes the Kozak sequence, a sequence that facilitates initial binding of mRNA to the small subunit of the ribosome and is required for initiation of translation. The PPKID4 sequence was amplified from the pAL₁-PPKID4 vector using the primers KID4-5B, which is complementary to the N-terminal portion of PPKID4 and includes the Kozak sequence and a BglII site upstream of
PPKID4 and KID4-3, which is complementary to the C-terminal portion of PPKID4 and contains the BamHI restriction site. KID4-5B: 5’TCAAGATCTGCCACCATGGGTCC GTCCACGCGACC3’, KID4-3: 5’TCAAGATCTGCCACCACCCGTTGCGTCCAGGTACAG3’. PCR samples contained 1 µl Gal4DBD-PPKID4 vector, 1 µM each of KID4-5B and KID4-3 primers, 300 µM dNTP’s in 1X ThermoPol buffer. Water was added to bring the final volume to 100µl. Immediately before the reaction was started, 1 µl Vent DNA polymerase was added to each sample. The PPKID4 insert was amplified using the following PCR program: 1) 94 °C for 2 minutes, 2) 94 °C for 30 seconds, 3) 72 °C for 30 seconds (-0.2 °C/cycle), 4) 72°C for 1 minute, 5) repeat steps 2-4 19x, 6) 72 °C for 3 minutes, 7) cool samples to 4 °C. The correct product was verified by separation on a Homo 20 PhastGel. Samples were purified using Qiagen PCR purification kit. The PPKID4 insert was digested with 40 units each of BglII and BamHI in a total volume of 30 µl for 15 hours at 37 °C. Digested DNA was purified using the QIAquick PCR purification kit. 10 µg of pAL1 vector was digested with 60 units each of BamHI and BglIII in a total volume of 30 µl for 4 hours at 37 °C. The digested vector was purified on a 1.2 % agarose gel using QIAquick gel extraction kit. Equal amounts of digested insert and vector were ligated using T4 DNA ligase at room temperature for 30 minutes. The ligation reaction (1 µl) was transformed into XL1 Blue electroporation competent E. coli cells and plated on LBKan agar. Plasmid DNA was prepared from five clones and sent for DNA sequencing. DNA from a correct pPPKID4 clone was prepared using a QIAfilter maxiprep kit to produce enough DNA for the reporter transcription assay. Resulting DNA was resuspended in TE buffer.
Cell culture – HEK293 cells

HEK293 cells (ATCC CRL-1573) were grown in DMEM complete media containing 10% FBS and 2mM L-glutamine at 37 °C and 5% CO₂. Cells were plated in 24-well plates for the transcription reporter assay or 60 mm culture dishes for western blotting 24 hours prior to transfection.

Transcription reporter assay

For 24-well plates, cells were transfected using SuperFect transfection reagent (Qiagen) with 800 ng of total DNA per well. Included were 400 ng of 5X Gal4 firefly luciferase reporter plasmid, 20 ng of a promoterless Renilla luciferase plasmid to normalize data for transfection efficiency, 5 ng of Gal4-peptide construct and 375 ng pBluescript SK+ carrier DNA. Where phosphorylation is indicated, 5 µM forskolin was added to cell media 30 hours post-transfection. Cells were harvested and assayed 36 hours after transfection.

Transfection

DNA was prepared, transferred to a sterile 5 ml tube and diluted with serum-free media. Superfect reagent was added to DNA and incubated at room temperature for 5-10 minutes. During incubation, media was removed from cells and the cells were washed with PBS. Serum-containing media was added to DNA, mixed and immediately added to cells. For specific amounts of DNA, Superfect reagent and media used, refer to Superfect Reagent handbook. The cells were incubated under normal growing conditions for 2 hours after which the media was removed, cells washed with PBS and fresh media added. Cells were incubated for 36 hours under normal growing conditions.
**Harvesting cells**

Media was removed and cells were washed once with PBS. 100 µl of 1X passive lysis buffer (Promega) was added to each well followed by incubation with shaking at room temperature for 15 minutes. Cell lysate was transferred to a tube and stored at –20 °C.

**Luciferase Assay**

The luciferase assay was performed using Promega’s Dual luciferase assay kit with a Turner Designs Model TD-20/20 single tube luminometer. This kit allows measurement of both firefly luciferase, produced from the Gal4 reporter gene, and *Renilla* luciferase, produced as a control, from the same sample. The luciferase assay reagent II (LARII) contains beetle luciferin, a substrate for firefly luciferase. The Stop & Glo Reagent quenches the firefly luminescence and contains coelenterazine, a substrate for *Renilla* luciferase. Cell lysate was thawed and brought to room temperature. Into a luminometer tube containing 100 µl of LARII, 20 µl of cell lysate was added and mixed. The luminescence from firefly luciferase was measured for 10 seconds after a 2 second delay. 100 µl of Stop & Glo reagent was added to the tube and the sample was mixed by vortexing. The luminescence from *Renilla* luciferase was measured for 10 seconds after a 2 second delay. The ratio of firefly luminescence to *Renilla* luminescence represents the normalized luciferase value.

**Western Blotting**

In 60 mm culture dishes, cells were transfected using SuperFect transfection reagent (Qiagen) with 4 µg of total DNA per well. Included were 2.5 µg of Gal4-peptide construct and 1.5 µg pBluescript SK+ carrier DNA. Cells were harvested 36-48 hours
after transfection. Media was removed and cells were washed with PBS. To lyse the cells, 350 µl 85° C 1X SDS-loading buffer was added to the dish and cells were scraped off the surface. The lysate was boiled for 10 minutes and homogenized using a QIAshredder column with centrifugation for 10 minutes at max speed in a microcentrifuge. The supernatant was transferred to a new tube and total protein concentration was estimated using a dot blot with a BSA standard. Approximately 30 µL of lysate was loaded per well (on average 40 µg of total protein) on a 16% SDS - polyacrylamide 10 cm x 8 cm minigel using 1 mm spacers. The gel was run at 40V on a Hoefer Mighty Small II apparatus (Amersham Biosciences). Protein was transferred to a PVDF membrane using a TE22 Mini Tank Transfer Unit (Amersham Biosciences) at 90V for 45 min. The PVDF membrane was blocked for 1 h in TBST containing 5% milk at room temperature. The membrane was then washed 3 x 5 min with TBST followed by incubation with anti-Gal4 DBD (1µg/mL) in TBST containing 3% milk for 1 h. The membrane was washed 3 x 5 min with TBST followed by incubation with anti-mouse IgG alkaline phosphatase conjugate (1:5000 dilution in TBST) for 30 min. The membrane was washed 3 x 5 min with TBST and 1 x 5 min with TBS. The membrane was then incubated with ECF substrate for 5 min and visualized on a Storm 840 Imager.

**Co-immunoprecipitation**

Cells in 60 mm dishes were washed with PBS and lysed with 450 µl RIPA buffer containing 1% PMSF, NaVO₃, and protease inhibitor. After incubation with gentle shaking for 15 min at 4 °C, cells were incubated on ice for 30 min. The cell debris was pelleted by centrifugation at 10,000 g for 10 minutes at 4 °C. The supernatant was incubated with 5µg anti-GAL4 DBD-agarose conjugate at 4 °C with mixing for 1 hour. The antibody and
bound proteins were pelleted by centrifugation at 500 g for 30 seconds at 4 °C. The pellet was washed twice with cold PBS and then resuspended in SDS-loading buffer. Samples were run on a 5% SDS-polyacrylamide gel. CBP was detected using the western blot protocol with the following changes: transfer protein to PVDF membrane for 2 hours at 80V, primary antibody used was anti-CBP and all buffers included 10 mM NaF as a phosphatase inhibitor.

**Pull-down Assay**

PPKID4\(^p\), PPKID6\(^U\), and KID-AB\(^p\) were synthesized as previously indicated without a c-terminal cysteine. HPLC purified peptides (1 mg) were coupled to 1 mL of Affi-Gel 10 for 4 hours at room temperature. Coupling was verified by decrease in protein absorbance at 280 nm. One 60 mm dish was lysed with 450 µl RIPA buffer containing protease and phosphatase inhibitors and clarified. Cell lysate (200 µl) obtained using immunoprecipitation protocol was incubated with peptide-coupled resin (400 µM final peptide concentration) for 2 hours at room temperature. Affi-Gel and bound proteins were washed with cold PBS + 0.1% Tween-20 and pelleted by centrifugation for 30 seconds at 500 g. The pellet was then resuspended in SDS-loading buffer and CBP was detected by western blotting.
References


