Inhibitors of Protein Kinases and Protein Phosphatases

Editors
Lorenzo A. Pinna and
Patricia T. W. Cohen

Springer

Handbook of Experimental Pharmacology

Volume 167

Inhibitors of Protein Kinases and Protein Phosphates

Contributors

D.R. Alexander, H.S. Andersen, N.R. Banner, R. Battistutta,
A.M. Berghuis, S.M. Blake, D. Bossemeyer, C. Breitenlechner,
D.L. Burk, A. Cheng, P. Cohen, S.W. Cowan-Jakob,
B.J. Druker, R. Engh, D. Fabbro, G. Fendrich, D.H. Fong,
P. Furet, M. Gaßel, J.D. Griffin, V. Guez, S. Herrero,
H. Hidaka, R.E. Honkanen, L.F. Iversen, C.B. Jeppesen,
S. Kumar, C. Kunick, C. Lampron, D.S. Lawrence, M. Leost,
O. Lozach, H. Lyster, P.W. Manley, L. Meijer, J. Mestan,
T. Meyer, N.P.H. Møller, S. Sarno, Y. Sasaki, S. Schmitt,
M. Shibuya, Y. Suzuki, M.L. Tremblay, N. Uetani,
A. Wakeling, M.H. Yacoub, G. Zanotti

Editors

Lorenzo A. Pinna and Patricia T.W. Cohen

Springer

Professor
Lorenzo A. Pinna
Università di Padova
Dipartimento di Chimica Biologica
Via Trieste 75
Padova 35121
Italy
e-mail: lorenzo.pinna@unipd.it

Patricia T. W. Cohen, Ph. D.
Professor of Molecular Biology,
MRC Protein Phosphorylation Unit,
School of Life Sciences
University of Dundee
MSI/WTB Complex, Dow St.
Dundee, DD1 5EH
Scotland, U. K.
e-mail: p.t.w.cohen@dundee.ac.uk

Preface

Nearly all aspects of cell life (and death) are controlled by the phosphorylation of proteins, which is catalysed by protein kinases (PKs) and reversed by protein phosphatases (PPs). The role of PKs can be likened to that of interpreters, who translate stimuli and signals into biochemical events. For this reason, PKs and PPs are themselves interlinked and highly regulated, forming complex communicative networks. Not surprisingly, therefore, the deregulation of PKs results in cell malfunction, eventually resulting in neoplastic growth and other diseases. This makes PKs attractive targets for drugs not only to combat cancer, but also for other global diseases, notably diabetes, inflammatory and infectious diseases, stroke, hypertension and Alzheimer's. Actually about half of all proto-oncogenes so far identified encode PKs, and oncogenesis frequently results from the activation and/or overexpression of PKs. For example, overexpression of the epidermal growth factor receptor tyrosine kinase is the cause of many cancers of epithelial cell origin. In other instances, however, the link of PKs with neoplasia is not so straightforward, and depends on defective interactions with cellular partners of PKs, susceptibility to particular metabolic conditions, abnormal levels of other regulatory components or the combination of several of these factors.

The attractiveness of PKs as targets is enhanced by the fact that they are enzymes, which are targetable molecules par excellence. Thus their biological activity can be turned off very easily and precisely by drugs that block the catalytic site. Virtually all PKs belong to the largest single family of enzymes, numbering over 500 and accounting for almost 2% of the proteins encoded by the human genome. They share similar catalytic domains that catalyse the transfer of phosphate from ATP to serine, threonine or tyrosine residues in key regulatory proteins. Nevertheless, the structures of the catalytic domains of PKs are sufficiently distinctive that it is possible to develop compounds that are highly selective for a particular PK. Even the highly conserved binding site for the substrate ATP is surrounded by structural elements with variable features that can be exploited for the design of specific inhibitors, and most of the PK inhibitors currently undergoing human clinical trials are of this type. Two PK inhibitors are already in clinical use for the treatment of cancers (Gleevec and Iressa), while another is the immunosuppressant of choice to prevent tissue rejection after organ transplantation (rapamycin). At least 30 other PK inhibitors are undergoing human clinical trials to treat cancers and other diseases. These have the potential to provide a significant impact

on the management of epithelial cancers, such as breast and lung cancer. The approval of Gleevec for the treatment of a form of leukaemia by the FDA in May 2001 and more recently for the treatment of stomach cancers was a landmark because it is the first drug to be developed by targeting specific PKs. Moreover, its spectacular clinical effects, with minimal side effects, have had an enormous impact on the pharmaceutical and biotechnology industry. As a result, PKs have become the second most important family of drug targets, 20%–30% of all drug development programmes now being concentrated in this area. Although most PK inhibitors currently under investigation as potential drugs are ATP site-directed ligands, the field is still in its infancy, and there is tremendous potential to develop different types of drugs that target the binding sites for the protein substrates or which prevent the activation of PKs, since many of these enzymes are arranged in 'cascades' in which one PK activates or inhibits another one. Longer-term strategies would involve approaches based on gene therapy in which the mutant PK would be replaced by the wild-type enzyme.

PPs have received less attention to date as potential drug targets than PKs. The empirical discovery of an immunosuppressant drug that revolutionised organ transplantation (ciclosporin) and the subsequent recognition that it is a specific inhibitor of one PP indicates that PPs can be effective drug targets. An anticancer agent also discovered empirically (fostriecin) is now recognised to be a PP inhibitor. Other PPs, such as PTP1B, are currently under active investigations as drug targets for the treatment of diabetes and other diseases. As with PKs, known PP inhibitors at present target the active site but since many PPs are complexes with regulatory subunits, there is a potential for developing drugs that target the binding site of these regulatory subunits or their interaction with regulators. Thus the expansion of PPs as suitable drug targets may eventually follow that of PKs.

This volume of HEP highlights the tremendous pharmacological potential of PK and PP inhibitors, by providing a thorough overview of the most remarkable achievements in the field and illustrating how beneficial these studies can be for the advancement of both basic knowledge on biological regulation and deregulation and for the clinical treatment of a wide spectrum of diseases.

List of Contributors

(Addresses stated at the beginning of respective chapters)

List of Contents

Part IV. Inhibitors in Clinical Use or Advanced Clinical Trials

Protein Kinase Inhibitors for the Treatment of Disease: The Promise and the Problems

P. Cohen

Medical Research Council Protein Phosphorylation Unit,
University of Dundee, MSI/WTB Complex, Dundee, Scotland, DD1 5EH, UK
p.cohen@dundee.ac.uk

1
The Promise

The reversible phosphorylation of proteins, catalysed by protein kinases and phosphatases, was first identified as a regulatory device in the 1950s, and it has been established for many years that this control mechanism regulates most aspects of cell life. However, it was only in the 1990s that interest in developing inhibitors of protein kinases and phosphatases started to enter centre stage (see Cohen 2002a,b for historical reviews). The first two drugs shown to target these classes of enzyme were cyclosporin, an inhibitor of protein phosphatase 2B (PP2B, also called calcineurin) (Liu et al. 1991) and rapamycin, an inhibitor of the protein kinase mTOR (mammalian target of rapamycin) (Heitman et al. 1991), which are the immunosuppressants that have permitted the widespread use of organ transplantation. However, these drugs were developed and approved for clinical use before their mechanism of action was identified. Fasudil, an isoquinoline sulphonamide that inhibits several protein kinases with relatively low potency, such as the Rho-dependent protein kinases (ROCK) (Davies et al. 2000), was developed by Hiroyoshi Hidaka in the 1980s and approved in Japan in 1995 for the treatment of cerebral vasospasm. ROCK can constrict blood vessels by inhibiting smooth muscle myosin phosphatase, but whether the clinical efficacy of fasudil results from its inhibition of ROCK, another protein kinase(s) or a completely different target, is unclear. Current information about this drug is discussed by Hidaka et al. (in Part 4).

Glivec (also called imatinib and STI-571), developed by Nick Lydon and his colleagues at Novartis, was the first drug to be developed by targeting a specific protein kinase and was approved for clinical use in the USA in 2001. It targets the protein tyrosine kinase c-Abl, which is mutated to the constitu-

tively active BCR-Abl fusion protein in nearly all cases of chronic myeloge-
nous leukaemia (CML). The spectacular efficacy and minimal side effects of
Glivec, first highlighted by Brian Druker, resulted in the most rapid approval
of a drug in FDA history and was a landmark event in this area. The devel-
opment of Glivec and its implications for the future of drug discovery in this
area are discussed by Fabbro et al. (in Part 4). Interestingly, Abl is not the
only protein tyrosine kinase targeted by Glivec. It also inhibits the c-Kit re-
ceptor tyrosine kinase and the platelet-derived growth factor (PDGF) recep-
tor. The c-Kit receptor is mutated to an abnormally active form in many gas-
trointestinal stromal tumours (GISTs) and the efficacy of Glivec for the treat-
ment GISTs is equally impressive, resulting in its approval for this therapeu-
tic use in 2002. The potential of Glivec to treat several types of cancer is dis-
cussed by Druker (in Part 4).

Following on from the successful launch of Glivec, Iressa a potent inhibi-
tor of the epidermal growth factor (EGF) receptor tyrosine kinase was ap-
proved in Japan in 2002 and in the USA in 2003 for the treatment of some
types of lung cancer. Developed by AstraZeneca, this drug is discussed by
Wakeling (in Part 4). Drugs that inhibit the vascular endothelial-growth fac-
tor (VEGF) or fibroblast growth factor (FGF) receptor tyrosine kinases are
undergoing phase III clinical trials and may be among the next protein ki-
nase inhibitors to be approved for clinical use. VEGF and FGF play key roles
in angiogenesis, and inhibitors of their receptors destroy the tumour's vas-
cular supply. For this reason these compounds may be useful for the treat-
ment of several types of cancer.

Compounds that inhibit protein serine/threonine kinases are also under-
going human clinical trials in a number of therapeutic areas. For example, at
least four companies have inhibitors of p38 mitogen-activated protein
(MAP) kinase in the clinic. These compounds suppress the production of tu-
mour necrosis factor (TNF) and some other proinflammatory cytokines and
show efficacy for the treatment of rheumatoid arthritis and other chronic in-
flammatory diseases. These programmes are discussed by Kumar and Blake
(in Part 2). In the same section, Meijer (in Part 2) discusses inhibitors of cy-
clin-dependent protein kinases (CDKs), which are undergoing clinical trials
as anti-cancer agents, and inhibitors of GSK3 which, although at the precli-
nical stage, have shown potential for the treatment of several diseases in-
cluding type II diabetes (Cline et al. 2002; Ring et al. 2003) and stroke (Cross
et al. 2001). Inhibitors of MAP kinase kinase 1 (MKK1, also called MEK)
and RAF (product of the proto-oncogene *Raf*) are undergoing clinical trials
as anti-cancer agents, and inhibitors of mixed lineage kinase 3 (MLK3) to
prevent neurodegeneration (reviewed in Cohen 2002b). However, this is only
the 'tip of the iceberg'. Over the past few years protein kinases have become
the second most studied group of drug targets after G protein-coupled re-
ceptors, accounting for a quarter or more of drug discovery programmes

worldwide. The number of protein kinase inhibitors undergoing human clinical trials at the present time almost certainly exceeds 100.

The discovery that PP2B, a serine/threonine-specific protein phosphatase, was inhibited specifically by cyclosporin highlighted the potential of protein phosphatases as drug targets, and programmes to develop specific inhibitors of several of these enzymes are underway. Protein tyrosine phosphatase IB (PTP1B) appears to be one of the enzymes that dephosphorylates and inactivates the insulin receptor, because mice that do not express it are hypersensitive to insulin and maintain normal blood glucose levels at half the normal circulating of insulin (Elchebly et al. 1999). In addition, these mice do not become obese when fed a high-fat, high-carbohydrate diet. For these reasons, PTPIB is potentially an attractive target for the development of a drug to treat diabetes and/or obesity, as discussed by Cheng et al. (in Part 3). However, although interesting compounds have been developed that are relatively specific inhibitors of PTP1B, as discussed by Møller (in Part 3), no inhibitors of this enzyme appear to have entered clinical trials. CD45 is another protein tyrosine phosphatase that is potentially an attractive drug target, because it is only expressed in cells of the immune system and is essential for T cell activation. Inhibitors of CD45 therefore have the potential to be effective immunosuppressants, but may lack the side effects associated with cyclosporin and rapamycin whose targets (PP2B and TOR) are expressed in nearly all cells and tissues. This topic is discussed by Alexander (in Part 3).

A number of toxins and tumour promoters are potent inhibitors of several members of one of the major classes of protein serine/threonine phosphatases, termed the PPP subfamily. They include the marine toxins responsible for diarrhetic seafood poisoning (okadaic acid and related compounds) and the algal toxins that are a threat to water supplies (microcystins) (reviewed in MacKintosh and MacKintosh 1994). Indeed, microcystins are the most potent liver carcinogens known to man. One might therefore predict that compounds which inhibit the catalytic subunits of these protein phosphatases would frequently be oncogenic and of little use as therapeutic agents. However, as discussed by Honkanen (Part 3), both fostriecin and cantharidin, which inhibit the same protein phosphatases, are cytotoxic for tumour cells and have been tested in phase I human clinical trials as anti-cancer agents. Not surprisingly, there are a number of side effects associated with the use of these compounds, and it seems more likely that drugs will eventually be developed that disrupt the functions of protein serine/threonine phosphatases in more subtle and specific ways. For example, the ability of the serine/threonine-specific protein phosphatase 1 (PP1) to dephosphorylate many proteins is controlled by its interaction with a great variety of 'targeting' subunits that direct it to specific subcellular locations and confer unique regulatory properties upon it. The form of PP1 associated with liver glycogen, which dephosphorylates and activates glycogen synthase, com-

prises the catalytic subunit of PP1 complexed to a glycogen-targeting subunit G_L. The ability of the PP1–G_L complex to dephosphorylate glycogen synthase is prevented when the active form of glycogen phosphorylase (termed phosphorylase a) binds to the extreme C-terminus of G_L, providing a mechanism for inhibiting glycogen synthesis when glycogenolysis is activated and vice versa (Armstrong et al. 1998). A drug that prevented the interaction of phosphorylase a with G_L would have the potential to lower the concentration of glucose in the blood by activating glycogen synthase and so stimulating the conversion of glucose into liver glycogen.

1.1
The Problems

There are over 500 protein kinases encoded by the human genome, most of which are members of the same superfamily. This has created a plethora of potential targets that can be studied in a unified way, but has highlighted the difficulty in developing compounds that are capable of inhibiting one of these enzymes specifically. The development of Glivec has shown that inhibition of more than one protein kinase can sometimes be beneficial, allowing the same drug to have more than one therapeutic use. However, more frequently one would expect such a lack of specificity to give rise to unwanted or unacceptable side effects. The recent availability of large panels of protein kinases (e.g. Davies et al. 2000; Bain et al. 2003) has been of considerable help in assessing the specificities of protein kinase inhibitors, and it is to be expected that such panels will continue to expand and eventually include the entire repertoire of protein kinases.

Lack of specificity may also mean that the therapeutic effect of a drug is actually mediated by inhibition of another protein kinase and not by inhibition of the kinase for which it was originally developed. For example, inhibitors of the cell cycle regulator CDK2 have been developed that suppress the proliferation of tumour cells, but these compounds may actually exert their therapeutic effects by inhibiting other protein kinases, such as CDK7 and/or CDK9, which are regulators of RNA polymerase II. It is therefore unclear whether the effects of these compounds are really mediated via CDK2. In order establish that the therapeutic effect of a drug is mediated by inhibition of a particular protein kinase one needs to show that the effects of the drug disappear in cells that express a drug-resistant mutant of the protein kinase (Eyers et al. 1999). It is possible to convert protein kinases to drug-resistant forms by single amino acid replacements (Brown et al. 1995; Eyers et al. 1998) so that, as for other types of drug, the development of drug resistance is a potential hazard. Mutations in Abl that make it resistant to Glivec are the cause of relapse in patients with chronic myelogenous leukaemia (Gorre et al. 2001). However, resistance to Glivec is mainly seen in patients

with the most advanced stage of this disease, where extensive genomic instability has already taken place.

Most of the protein kinase inhibitors developed thus far target the ATP-binding site and must therefore be of sufficient potency to compete with the millimolar concentrations of ATP that are present in the intracellular milieu. Clearly, it is possible to develop compounds with the requisite in vivo potency, as shown by the number of compounds undergoing human clinical trials. However, this remains a challenging problem, especially for protein kinases that bind ATP particularly tightly. Some of the most interesting protein kinase inhibitors developed thus far, including Glivec (Schindler et al. 2000) and the p38 MAP kinase inhibitor BIRB 796 (Pargellis et al. 2002), not only target the ATP-binding site, but also trigger structural changes that induce the inactive conformations of these protein kinases. Two other compounds, PD 98059 and U0126, do not target the ATP-binding site at all, but bind to the inactive conformation of MKK1, preventing it from being activated by the protein kinase Raf (Alessi et al. 1995; Davies et al. 2000). The development of more compounds that prevent one protein kinase from activating another may be a promising strategy for novel drug development in this area, since many of these enzymes are components of protein kinase 'cascades'. Another way of generating compounds that are not ATP-competitive would be to target the binding sites for protein substrates, a topic discussed by Lawrence (in Part 1).

There are about 150 protein phosphatase catalytic subunits encoded by the human genome, and they fall into three main superfamilies. The generation of compounds that discriminate between different protein phosphatases is therefore also a challenging one. However, in contrast to protein kinases, the option of targeting an ATP binding pocket does not exist. Moreover, the protein substrate-binding cleft can be very polar, as in the case of PTP1B (Kellie 2003). This has made it difficult to develop compounds that combine high potency with cell permeability. The only protein phosphatase inhibitor that has advanced to human clinical trials, cyclosporin, inhibits PP2B in an unusual way; it binds to the protein cyclophilin, and the cyclosporin–cyclophilin complex then inhibits the protein phosphatase (Liu et al. 1991). As discussed earlier, it seems more likely that the future of drug discovery in this area may lie in targeting the regulatory subunits of serine/threonine-specific protein phosphatases.

Finally, it is important to mention that inhibitors of protein kinases are not only becoming important for the treatment of disease, but also as reagents for the study of cell signalling. The huge number of citations garnered by the publications that have introduced these compounds to the scientific community are a reflection of the widespread need for these compounds by the scientific community. For example, I was surprised to learn from the Institute for Scientific Information that the paper we published in 1995 with David Dudley and Alan Saltiel at Parke Davis on the mechanism

of action of PD 98059 (Alessi et al. 1995) was the UK's most frequently cited original research paper over the past 10 years in the fields of biology and biochemistry, while our publication with Peter Young and John Lee at SmithKline Beecham on the specificity of SB 203580 (Cuenda et al. 1995), a prototypic p38 MAP kinase inhibitor, was the UK's sixth most cited original research paper over this period. Although many compounds are advertised for sale as 'specific protein kinase inhibitors', in practice many have turned out to inhibit so many protein kinases that conclusions drawn from their use are likely to be erroneous (Davies et al. 2000; Bain et al. 2003). The number of really useful protein kinase inhibitors that are available commercially is still rather limited, but the number will increase considerably over the next few years. I believe that pharmaceutical companies have much to gain from the discoveries that will be made by exploiting these compounds, and it is to be hoped that many more will be released for general use in the future.

References

Alessi DR, Cuenda A, Cohen P, Dudley DT, Saltiel AR (1995) PD98059 is a specific inhibitor of the activation of mitogen-activated protein kinase kinase in vitro and in vivo. J Biol Chem 270:27489–27494

Armstrong CG, Doherty MJ, Cohen PTW (1998) Identification of the separate domains in the hepatic glycogen-targeting subunit of protein phosphatase 1 that interact with phosphorylase a, glycogen and protein phosphatase 1. Biochem J 336:699–704

Bain J, McLauchlan H, Elliott M, Cohen P (2003) The specificities of protein kinase inhibitors: an update. Biochem J 371:199–204

Brown EJ, Beal PA, Keith CT, Chen J, Shin TB, Schreiber SL (1995) Control of p70S6 kinase by kinase activity of FRAP in vivo. Nature 377:441–446

Cline GW, Johnson K, Regittnig W, Perret P, Tozzo E, Xiao L, Damico C, Schulman GI (2002) Effects of a novel glycogen synthase kinase-3 inhibitor on insulin-stimulated glucose metabolism in Zucker diabetic fatty (fa/fa) rats. Diabetes 51:2903–2910

Cohen P (2002a) The origins of protein phosphorylation. Nat Cell Biol 4:E127–E130

Cohen P (2002b) Protein kinases—the major drug targets of the twenty-first century? Nat Rev Drug Discov 1:309–315

Cross DA, Culbert AA, Chalmers KA, Facci L, Skaper SD, Reith AD (2001) Selective small molecule inhibitors of glycogen synthase kinase-3 activity protect primary neurons from death. J Neurochem 77:94–102

Cuenda A, Rouse J, Doza YN, Meier R, Cohen P, Gallagher TF, Young PR, Lee JC (1995) SB 203580 is a specific inhibitor of a MAP kinase homologue which is stimulated by cellular stress and interleukin-1. FEBS Lett 364:229–233

Davies SP, Reddy H, Caivano M, Cohen P (2000) Specificity and mechanism of action of some commonly used protein kinase inhibitors. Biochem J 351:95–105

Elchebly M, Payette P, Michaliszyn E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML, Kennedy BP (1999) Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase 1B gene. Science 282:1544–1548

Eyers PA, Craxton M, Morrice N, Cohen P, Goedert M (1998) Conversion of SB-203580-insensitive MAP kinase family members to drug-sensitive forms by a single amino acid substitution. Chem Biol 5:321–328

Eyers PA, Van den Ijssel P, Quinlan R, Goedert M, Cohen P (1999) Use of a drug resistant mutant of stress activated protein kinase 2a/p38 to validate the in vivo specificity of SB203580. FEBS Lett 451: 191–196

Gorre ME, Mohammed M, Ellwood K, Hsu N, Pacquette R, Rao PN, Sawyers CL (2001) Clinical resistance of STI 571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 293:876–880

Heitman J, Mowa NR, Hall MN (1991) Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science 253:905–909

Kellie S (2003) Protein tyrosine phosphatases: potential roles in disease. In: Watling KJ (ed) Celltransmissions. Sigma-Aldrich, St Louis, pp 3–8

Liu J, Farmer JD, Lane WS, Friedman J, Weissman I, Schreiber SL (1991) Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell 66:807–815

MacKintosh C, MacKintosh RW (1994) Inhibitors of protein kinases and phosphatases. Trends Biochem Sci 19:444–447

Pargellis C, Tong L, Churchill L, Cirillo PF, Gilmore T, Graham AG, Grob PM, Hickey ER, Moss N, Pav S, Regan J (2002) Inhibition of p38 MAP kinase by utilising a novel allosteric binding site. Nat Struct Biol 9:268–272

Ring DB, Johnson KW, Henriksen EJ, Nuss JM, Goff D, Kinnick TR, Ma ST, Reeder JW, Samuels I, Slabiak T, Wagman AS, Hammond M-EW, Harrison SD (2003) Selective glycogen synthase kinase 3 inhibitors potentiate insulin activation of glucose transport and utilisation in vitro and in vivo. Diabetes 52:588–595

Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J (2000) Structural mechanism for STI 571 inhibition of Abelson tyrosine kinase. Science 289:1938–1942

Part I
General Aspects of PKs Inhibition

New Design Strategies for Ligands That Target Protein Kinase-Mediated Protein–Protein Interactions

D. S. Lawrence

Department of Biochemistry, The Albert Einstein College of Medicine,
1300 Morris Park Ave, Bronx, NY 10461, USA
lawrence@medusa.bioc.aecom.yu.edu

Abstract Protein–protein interactions serve as the molecular engine that drives the formation and disassembly of intracellular signaling pathways. Antagonists of these interactions could play key roles as both biological reagents and therapeutic compounds. However, much of the early work in this area with peptides revealed that these species, in general, bind with modest affinity to their protein targets. In addition, when these studies first commenced nearly 20 years ago, the technology for the intracellular delivery of peptides and modified analogs thereof was rudimentary. In the intervening years, not only has this technology dramatically improved, but the global role that protein–protein interactions play in transducing intracellular signals has become simply too obvious to ignore. With the introduction of combinatorial library methods, it is now a simple matter to identify consensus sequences recognized by protein interaction domains. An array of strategies has now been developed to transform these otherwise modest binding consensus sequences into high-affinity ligands. These strategies include the design of high-affinity replacements for key amino acid residues in consensus peptides, the construction of

multidomain-binding peptides, and the structural modification of consensus sequence peptides. In several of these instances, unprecedented affinity (<nM) and selectivity (>1,000-fold versus closely related protein targets) have been achieved.

1
Introduction

Protein–protein interactions serve as the adhesive that drives the assembly of signaling pathways. However, this adhesive is transient in nature. Once the cell has acknowledged the environmental stimulus, signaling pathways must rapidly disassemble to restore the cell to its resting state. At first glance, agents that selectively target key protein–protein interactions would appear to serve as ideal inhibitors of cell signaling as well as potential therapeutics. First, protein–protein interactions are typically exemplified by well-defined consensus sequences, which can often be reasonably selective for a given protein–protein pair. Consequently, the preparation of inhibitors of protein–protein interactions appears, at least on paper, to be reasonably straightforward since, the acquisition of preferred consensus sequences employs simple and well-defined methods. Second, the intracellular levels of protein–protein-binding partners rarely surpass low micromolar amounts, thereby rendering competition with endogenous substrates relatively unimportant. In spite of these apparent advantages, the overwhelming majority of reported protein kinase inhibitors target the ATP-binding site, a region common to all protein kinases, non-protein kinases, and many other ATP-binding proteins. Furthermore, the intracellular concentration of ATP (~1–10 mM) is much larger than its K_m (serine/threonine kinases ~1–10 µM; tyrosine kinases ~20–50 µM), which all but assures that the ATP-binding site will be saturated with ATP. The consequence of the latter is that inhibitors that target the ATP-binding site must be present at intracellular concentrations that significantly exceed their in vitro-determined K_i values. Finally, the acquisition of ATP analogs that specifically target individual protein kinases requires the initial screening of a large starting library of potential inhibitor candidates. This is then followed by a substantial synthetic effort that involves the preparation of secondary and tertiary libraries based on initially identified leads. The notion of disrupting signaling pathways via antagonists of protein–protein interactions has been unpopular for a number of reasons, including issues related to potency, intracellular stability and uptake, and general bioavailability (i.e., with respect to therapeutics). However, recent advances in various delivery technologies coupled with our in-

creasing understanding of the widespread participation of protein-binding domains in signaling, has led to a renewed interest in the development of anti-signaling agents that disrupt intracellular protein–protein interactions.

Given the long-dormant state of this field, which is characterized by a recent reawakening, a broad overview of the general area of protein kinase-mediated protein–protein interactions and their corresponding antagonists is provided. This includes a summary of the methods employed to obtain consensus sequence information, a general synopsis of protein-binding domains, and finally a description of antagonists of protein–protein interactions as well as emerging strategies to acquire ever more potent and selective inhibitory agents.

2
Identification of Consensus Sequences

2.1
Degradation of Protein Ligands

Amino acid recognition sequences that drive protein–protein interactions were initially identified via partial digestion of one of the protein-binding partners. Fragments that were determined to retain binding potency were then sequenced. Further refinement of the amino acid recognition sequence could then be explored via the preparation of synthetic peptides. This strategy is best exemplified by the work described in the 1980s on the potent "heat-stable" inhibitor of the cyclic adenosine monophosphate (cAMP)-dependent protein kinase (PKA) known as PKI (protein kinase inhibitor). Krebs, Walsh, and their colleagues (Scott et al. 1985a,b; Cheng et al. 1986; Scott et al. 1986; Van Patten et al. 1986; Glass et al. 1989) identified a series of peptides that serve as extraordinarily potent inhibitors (K_i<50 nM) of PKA. Protease digestion of the isolated protein furnished a 20-mer peptide that acts as a competitive inhibitor versus peptide substrate with a K_i in the subnanomolar range. These investigators demonstrated that the sequence Gly-Arg-Thr-Gly-Arg-Arg-Asn-Ala-Ile is the active site-directed component of PKI, where the Ala residue is positioned at the site normally reserved for the phosphorylatable serine. Indeed, subsequent studies demonstrated that insertion of serine in active site-directed sequences derived from PKI generates powerful peptide substrates (Mitchell et al. 1995). However, the new library-based methods introduced in the 1990s have largely supplanted the biochemical approaches for identifying amino acid sequences recognized by protein interaction domains. The new methodologies are not only significantly less labor intensive than their classical counterparts, but are also able to bypass the need for large quantities of both binding partners (for digestion and sequencing purposes).

2.2
Synthetic Peptide Libraries

A large number of different library strategies using synthetic peptides have been described. These approaches include one-bead/one-peptide libraries (Wu et al. 1994), solution mixtures of peptides (Songyang et al. 1994), one-well/one-peptide strategies (Lee and Lawrence 1999), peptides on chips (Houseman et al. 2002), and even proteins on chips (Zhu et al. 2000). A detailed description of the vast array of peptide library strategies now available is well beyond the scope of this review. However, all of these methods offer a rapid means to quickly identify preferred amino acid sequences in what is typically a single experiment. Peptide-based libraries also permit the use of amino acid derivatives beyond the standard genetically encoded residues (e.g., post-translationally modified residues such as phosphoTyr, hydroxyPro, etc.). In addition, many of the methods not only identify a preferred consensus sequence, but also often furnish an assessment of the range of residues permitted at a given position on the peptide ligand.

Each of the peptide library strategies enjoys certain advantages while enduring specific disadvantages:

1. One-bead/one-peptide libraries are extremely easy to prepare via split-and-pool synthesis (Lam et al. 2003). However, these libraries are commonly composed of a mixture of millions of beads, with each bead possessing a unique peptide sequence. Consequently, a screening method must be devised so that the bead containing the tightest binding ligand can be readily identified. Possibilities include the use of a target protein that contains an appended fluorophore or is conjugated to an enzymatic reporter. Beads can also be identified via the introduction of radioactivity (i.e., the use of $[\gamma\text{-}^{32}P]ATP$). Once leads have been identified, the beads are isolated and the bound peptides identified by microsequencing. Given the heavy reliance upon the latter, the use of uncommon hypermodified residues is severely restricted.

2. Soluble peptide library mixtures have also been utilized to identify consensus sequences (Songyang and Cantley 1998). These libraries are prepared by treating the growing peptide chain with a mixture of the standard amino acid derivatives. The actual ratio of the amino acids introduced during the coupling reaction is based upon the relative coupling efficiencies of the individual residues. Consequently, a particular residue that couples sluggishly (e.g., Arg) is present at a greater relative ratio than one that couples readily (e.g., Gly). Following completion of the synthesis, the peptide mixture is cleaved from the resin and subsequently employed for consensus sequence identification. The latter is achieved by selective enrichment of the binding sequence, often using an affinity column. For example, protein kinase-catalyzed phosphorylation of the mixture is allowed to proceed until a small fraction (<1%) of the total peptide is phosphorylated. The phosphopeptide

mixture is subsequently isolated and sequenced as a mixture. Each position on the peptide is not identified as a single residue, but rather as the relative abundance of all the amino acid residues at a particular site. The residue present in the largest amount at a given position is taken as the one most favored at that site. However, since a peptide mixture, as opposed to a single peptide, is sequenced, this strategy does not yield sequences of unique peptides but merely determines the preferences for particular residues at specific positions. An inherent assumption of this method is that selection at each position is independent of the adjacent amino acids. Consequently, this technique ignores the possibility that two or more residues can act in a synergistic fashion to promote target protein affinity.

3. The one-well/one-peptide approach ("parallel synthesis") (Granier 2002) employs pure peptides that are spatially segregated from one another (Lee and Lawrence 1999). This technique has the advantage that the sequence of each peptide in each well is verified in advance. Furthermore, a wide assortment of hypermodified amino acid residues can be employed, since the synthesis history of each peptide in each well in known. An obvious disadvantage is that the size of these libraries, by necessity, is much smaller than those described in points *1* and *2* above. Variations that employ spatially segregated mixtures ("positional scanning") have been reported that address this concern (Houghten et al. 1996).

4. Peptide chips represent the solid phase version of the method described in *3* (Houseman et al. 2002). The added advantage of this system is the higher spatial density, and therefore smaller chip size [membranes have been employed in this technique as well (Frank 2002)]. However, the increased spatial density of the individual peptide "colonies" can come at a cost. Although methods that employ fluorescence detection of target protein binding will work well in this system, other common methods, such as those that utilize radioactivity, cannot be applied to ultra high-density chips.

2.3
Phage Display

Phage display is a genetically encoded peptide library strategy (Scott and Smith 1990; Smith and Scott 1993). In brief, peptides are displayed on the capsid protein of filamentous phage. Each virion particle displays a unique peptide sequence on its surface. Millions of phage clones are exposed to the protein target of interest and the affinity purified particles then amplified in *Escherichia coli*. Subsequent rounds of selection furnish a few "lead" clones, from which the displayed sequences can be determined via sequencing of the viral DNA coding region. Phage display has been used to acquire peptide ligand sequences for a wide variety of protein interaction domains (Smothers et al. 2002). The obvious limitation here is that the genetic basis for this method restricts the range of amino acids to the 20 standard residues.

3
The Protein-Binding Domains of Protein Kinases

Protein kinases are, first and foremost, catalysts that promote the transfer of a phosphoryl group from ATP to the acceptor hydroxyl moiety of serine, threonine, and/or tyrosine. The serine, threonine, and tyrosine residues must be embedded within the proper amino acid sequence in order to be recognized by a given protein kinase, a fact exemplified by the large number of synthetic peptide-based substrates that have been devised for scores of protein kinases. Of all the protein-binding domains contained within protein kinases, the active site region displays the greatest diversity in terms of sequence recognition. However, since all members of the protein kinase family utilize the same phosphoryl donor (ATP) and acceptors (serine, threonine, tyrosine), it is perhaps not too surprising that the conformation of the active site region is remarkably well-conserved (Johnson et al. 1998; Huse and Kuriyan 2002). The "protein kinase fold" is composed of two separate lobes, commonly designated as the N- and C-terminal lobes. The former is the smaller of the two and is composed of five antiparallel β-strands and a single α-helix. The larger C-terminal lobe is primarily α-helical in structure. ATP resides in a cleft that lies at the interface between the N- and C-terminal lobes. By contrast, the peptide/protein phosphoryl acceptor is primarily associated with the C-terminal lobe. The catalytic domain of protein kinases can assume active and inactive conformational states. The lobes in the former migrate toward one another, thereby closing the active site and promoting catalysis.

Protein kinases are commonly differentiated on the basis of their preferred phosphoryl acceptor group on the protein substrate: either the aliphatic hydroxyl moieties of serine and threonine ("serine/threonine protein kinases") or the aromatic phenol of the tyrosine residue ("tyrosine protein kinases"). A few protein kinases display the property of "dual specificity" in terms of their ability to recognize and phosphorylate both aliphatic and aromatic alcohols on peptides or proteins in vitro (fewer still display this property in living cells) (Dhanasekaran and Premkumar Reddy 1998; Marin et al. 1999). However, in a very strict sense, the segregation of protein kinases into these separate camps most likely has less to do with the protein kinases themselves and more to do with the fact that the genetic code is limited to only 20 different amino acids. For example, PKA, a well-known serine/threonine-specific protein kinase, phosphorylates appropriately designed aromatic alcohols (e.g., **1–3**) (Lee et al. 1994).

OH

HO   OH

peptide-NH   peptide-NH   peptide-NH

1   2   3

OH   (CH$_2$)$_4$OH   (CH$_2$)$_{10}$OH

peptide-NH   peptide-NH   peptide-NH

4   5   6
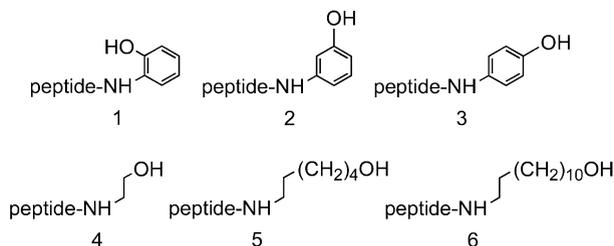
Furthermore, Src, an equally well-established tyrosine-specific protein kinase, phosphorylates aliphatic alcohols (e.g., **4–6**) (Lee et al. 1995a,b). Obviously, from the biological point of view, compounds **1–6** are mere curiosities. However, in terms of designing sensors, substrates, and antagonists of protein kinase-mediated protein–protein interactions, derivatives **1–6** are an important reminder that, unlike cells, chemists are not limited to the standard 20 amino acids fixed by the genetic code. Indeed, early work with consensus sequence peptides *containing conventional amino acids* is, in large part, responsible for the prevailing notion that inhibiting signaling pathways via disruption of protein–protein interactions is a strategy doomed to failure.

In addition to the "active site specificity" (i.e., serine/threonine versus tyrosine) of protein kinases, these enzymes display a preference for the amino acid sequence that encompasses the phosphorylatable residue (the "sequence specificity"). Compilations of sequences phosphorylated by protein kinases are available and these will not be recapitulated here (Pinna and Ruzzene 1996). However, certain trends are apparent:

1. The overwhelming majority of protein kinases will also phosphorylate simple peptides, thereby rendering the in vitro assay of these enzymes fairly straightforward. In addition, this demonstrates that the protein-binding region of these enzymes is sufficiently structurally well established to recognize substrates on its own (i.e., large intact protein substrates are not required for the protein kinase to assume an active state). However, the notion that the protein-binding region is the only site on the protein kinase that is responsible for substrate specificity is decidedly untrue.

2. Absolute protein kinase specificity is not encoded within the substrate-binding site. In other words, the consensus sequence surrounding the phosphorylatable residue, although an important parameter of protein kinase recognition, is not the sole determinant of specificity. For example, the cyclic guanosine monophosphate (cGMP)-dependent protein kinase (PKG) and PKA, members of the same protein kinase subfamily, display largely overlapping specificities with respect to simple peptide substrates (Mitchell et al. 1995; Wood et al. 1996). Certain protein kinase C (PKC) isoforms likewise phosphorylate the same peptides as PKA and PKG; however, the sequence preference of PKC is broad enough that PKC peptide substrates have

been devised that are recognized by neither PKA nor PKG (Yan et al. 2000). In short, it is unlikely that a peptide composed of only conventional amino acids will serve as an absolutely specific substrate for any given protein kinase. Indeed, Cohen and his colleagues have used this notion to generate a small set of peptides that serve as general substrates for more than three dozen different protein kinases (Ross et al. 2002).

3. Consensus sequence-containing active site-directed peptides are generally poor inhibitors. This fact is responsible for much of the common belief that targeting the substrate recognition site in particular, and protein–protein interactions in general, is an untenable strategy. Perhaps the best-known example is kemptide, Leu-Arg-Arg-Ala-Ser-Leu-Gly, which serves as an excellent substrate for PKA. The K_m for this peptide is less than 20 μM, a value that was (incorrectly) taken as a reflection of the binding constant of the peptide. However, the corresponding nonphosphorylatable peptide, Leu-Arg-Arg-Ala-Ala-Leu-Gly, is an exceedingly poor PKA inhibitor (K_i=320 μM) (Whitehouse et al. 1983). Much of the early discussion concerning the ineffectiveness of the Ala-substituted peptide centered on the possibility that the serine hydroxyl group (missing in the inhibitor) promotes binding affinity by two orders of magnitude. However, subsequent detailed enzymological studies revealed that the K_m value is a complex parameter that is dependent upon more than just the microscopic rate constants that control the active site association and dissociation of peptide substrate (Adams and Taylor 1992). In an analogous vein, poor inhibitors of tyrosine kinases (in which the phosphorylatable tyrosine residue was replaced with a phenylalanine) have been noted. Nevertheless, a few exceptions to the "rule" that active site-directed peptides serve as poor inhibitors are known. A naturally occurring "heat stable" protein-based inhibitor (PKI) of PKA is a powerful inhibitor (K_i<1 nM) (Whitehouse and Walsh 1983). A peptide fragment of PKI, most notably Gly-Arg-Thr-Gly-Arg-Arg-Asn-Ala-Leu-Gly-amide (where the Ala represents the site that would be reserved for serine) was reported to have a K_i of 36 nM (Glass et al. 1989). As an aside, the K_i for this peptide was originally determined under conditions of low salt; when the inhibitory potency of this peptide was subsequently reexamined several years later under more physiologically conditions, the K_i was found to be 500 nM (Wood et al. 1998). Nevertheless, the latter value does suggest that it is feasible to devise reasonably potent inhibitors based on standard amino acid residues alone. Indeed, perhaps the most outstanding example of this is the extraordinarily potent 24 amino acid-containing PKG selective inhibitor Arg-Gln-Ile-Lys-Ile-Trp-Phe-Gln-Asn-Arg-Arg-Met-Lys-Typ-Lys-Lys-Leu-Arg-Lys-Lys-Lys-Lys-Lys-His (Dostmann et al. 2000). This peptide is likely engaged in interactions beyond the immediate vicinity of the active site. Indeed, there appears to be a general consensus of opinion that, unlike targeting the ATP-binding site, effective inhibitors of protein–protein interactions must coordinate to a relatively large surface area.

In addition to the active site, there are several other protein interaction domains that are commonly affixed to protein kinases. These include the SH2 and SH3 domains, which are prevalent among the tyrosine protein kinases. PDZ, LIM, WW, PTB, and others are found in many protein kinases and/or in the adaptor proteins that help to transduce the activity of these enzymes. In addition, protein kinases themselves serve as ligands for protein interaction domains present on anchoring proteins. Indeed, peptide-derived inhibitors that bind to these anchoring proteins and thereby block protein kinase docking, have been described (Csukai and Mochly-Rosen 1999). The primary focus of this chapter is on the acquisition of active site-directed peptide-based inhibitors and the emerging strategies to acquire ever more potent and selective agents.

4
Strategies for the Acquisition of Potent and Selective Peptide-Based Inhibitors of Protein Kinases

Although a few exceptions are known, in general, conventional peptides display modest affinities and poor selectivities for the protein interaction domains contained within protein kinases. Biological systems appear to have little need for high-affinity ligands for active sites, SH2, SH3, LIM, PDZ, and other protein-interaction domains due to the transient nature of signaling pathways. However, it is abundantly clear that biological systems have mastered the issue of selectivity. Selective expression of only certain protein kinases in specific cell types, or at precise intervals during the lifetime of the cell, offers one means to navigate the tricky waters of intracellular selectivity. Spatial segregation of protein kinases to specific intracellular sites represents another means by which selectivity can be achieved. Finally, given the comparatively large size of these proteins, and their correspondingly well-defined structures, selectivity may simply be attained via a highly precise three-dimensional choreography of interactions between binding partners. Consequently, the design of potent and selective artificial antagonists of protein–protein interactions represents a significant challenge, albeit an exciting one. The primary advantage enjoyed by the chemist is that he or she is not restricted to the 20 standard amino acids designated by the genetic code.

The acquisition of agents that target protein–protein interaction sites has the potential to be relatively straightforward. Consensus sequences are easy to identify. However, it is necessary to develop the tools and/or strategies that can convert peptides containing these sequences into agents that recapitulate the high selectivities observed in biochemical pathways while significantly surpassing the affinities that intracellular binding partners display for one another. The tools and strategies to achieve the twin goals

of potency and selectivity, within the framework of relatively small ligands (cf., proteins), are outlined below in three separate, but interrelated sections.

4.1
Mimetics of Key Residues in Consensus Sequence Peptides

4.1.1
Serine Analogs

As noted above, one of the first indications that the acquisition of effective peptide-based inhibitors for protein kinases might be problematic was the replacement of the phosphorylatable serine residue in the PKA substrate kemptide with an alanine to create a dead-end inhibitor. The latter proved to be an unexpectedly weak inhibitory agent (K_i>300 μM versus the K_m for kemptide <20 μM). One of the explanations offered for the low affinity, namely loss of the hydroxyl serine side chain as a potential hydrogen bond donor, was subsequently shown to be incorrect. However, this notion does suggest that there may be ways to improve upon the use of alanine as a non-phosphorylatable replacement for serine.

Coward and his colleagues were the first to suggest that the phosphorylatable residue in an active site-directed peptide could be substituted with an analog that is able to also engage the ATP-binding site (i.e., a bisubstrate inhibitor) (Lashmet et al. 1983).
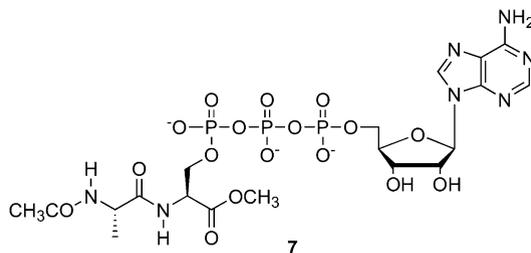
Although the ATP-γ-Ala-Ser ester **7** does not possess the requisite peptide framework for it to serve as a protein kinase inhibitor, it is a model of the type of compounds that were eventually prepared more than a decade later.

Gibson and his colleagues were the first to report the synthesis of adenosine phosphopeptides in a solid phase format (Medzihradszky et al. 1994).
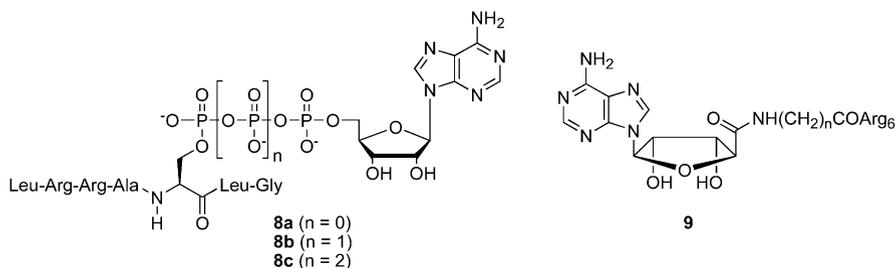
These investigators described the preparation and characterization of several analogs (**8**) of kemptide. Although the IC$_{50}$ values of **8a**, **8b**, and **8c** (935 μM, 226 μM, and 68 μM, respectively) are modest, these inhibitors are significantly more effective than the simple Ala-containing analog Leu-Arg-Arg-Ala-Ala-Leu-Glu. The authors found that **8** displays a competitive inhibition pattern versus variable ATP, but such a pattern was not observed with respect to variable peptide substrate. One might expect that a bisubstrate analog would exhibit competitive patterns versus both ATP and phosphorylatable peptide. However, the absence of double competitive behavior does not necessarily rule out the two-site binding model. Strictly speaking, competitive behavior is observed for an inhibitor only if that inhibitor and the corresponding substrate bind in a mutually exclusive fashion to the *same enzyme form*. PKA is known to exhibit a primarily ordered mechanism with ATP binding first (Whitehouse et al. 1983). Consequently, one would expect compound **8** and ATP to associate with the same enzyme form, namely the free enzyme, and thereby exhibit competitive behavior. By contrast, given the nature of the ordered mechanism, peptide substrate preferentially coordinates to the enzyme–ATP complex, which would thereby rule out a competitive pattern with **8**.

Recently Uri and his colleagues have described a series of bisubstrate analogs that dispenses with the phosphoric anhydride portion of the ATP moiety (Loog et al. 1999). These investigators employed an adenosine-5'-carboxylic acid derivative as the ATP mimic which, using a variety of linkers, was appended to the N-terminus of an arginine rich peptide (**9**). The most effective analogs displayed IC$_{50}$ values of between 100 and 300 nM for PKA and PKC. These bisubstrate analogs have been used to affinity-purify protein kinases (Loog et al. 2000). In addition, membrane-permeable fluorophore-labeled bisubstrate derivatives have been prepared (Uri et al. 2002; Viht et al. 2003).

In a departure from the ATP-based bisubstrate strategy, Sergheraert and colleagues designed (ATP mimics)-linker-substrate analogs (Ricouart et al. 1991). Isoquinoline and naphthalene sulfonic acid derivatives served as ATP replacements. The most potent of the several derivatives prepared was compound **10**.

10

11
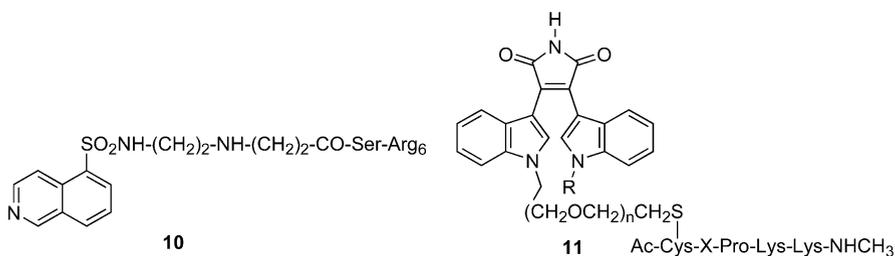
The latter exhibits a 25-fold selectivity in favor of PKA (K_i=4 nM) versus PKC (K_i=100 nM). The inhibitory potency of this derivative is impressive when one considers the fact that the ATP mimic alone is a nearly three orders of magnitude poorer inhibitor than **10**. However, since the ATP analog is appended off the N-terminus of the peptide, an unanswered question is the nature of the requisite structural requirements to replace a serine moiety that is contained within the interior of a consensus sequence. Finally, **10** acts as a competitive inhibitor versus variable ATP, but is not competitive with respect to variable peptide substrate. Sasaki, Maeda, and their coworkers likewise utilized an ATP analog (a bisindolylmaleimide) to prepare a series of bisubstrate inhibitors **11** that are designed to target the cyclin-dependent protein kinase, cdc2 (Sasaki et al. 1998). The best inhibitors display IC$_{50}$ values in the low micromolar range (where X=no amino acid). However, when X=Ser, the inhibitory potency is reduced by two orders of magnitude.

4.1.2
Tyrosine Analogs

In an analogous vein to serine/threonine protein kinases, peptide-based inhibitors of tyrosine kinases were initially prepared by substituting the phosphorylatable tyrosine with the nonphosphorylatable phenylalanine. In general, the phenylalanine-for-tyrosine replacement generates exceedingly poor inhibitory agents (K_i>1 mM).

12   **13**   **14**
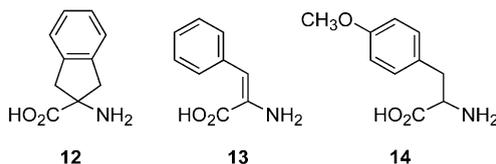
Several peptides containing tyrosine analogs (**12–14**) were reported in the 1980s, but these derivatives proved to be ineffective as inhibitors (Wong and Goldberg 1984; Shoelson et al. 1989).

Graves and his colleagues described the first example of an effective peptide-based tyrosine kinase inhibitor (Yuan et al. 1990).

15   **16**
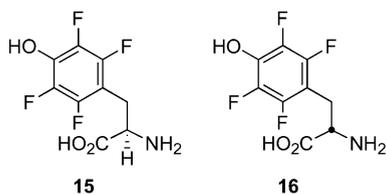
   The inhibitory agent, which contains the tetrafluorotyrosine moiety **15**, targets the insulin receptor with a K_i of 4 μM. The rationale for the use of the fluorinated tyrosine analog was based on the presumed mechanism of catalysis. These investigators reasoned that an active site base partially removes the aromatic hydroxyl proton during the transition state of the enzyme-catalyzed phosphoryl transfer reaction from ATP to the acceptor phenol. Presumably, the enzyme stabilizes this partial-negative charge on the phenol/phenoxide during the transition state, which suggests that a tyrosine analog that is negatively charged might be well accommodated within the active site. The four fluorine substituents not only lower the pK_a of the phenol, thereby promoting ionization to the phenoxide at physiological pH, but in addition they render the phenoxide less nucleophilic than its natural counterpart. These investigators also prepared the corresponding d-analog **16**, which also displays promising inhibitory activity (K_i=20 μM). Interestingly, although both **15** and **16** serve as competitive inhibitors versus variable peptide substrate, the l-analog directly competes with ATP as well, whereas the d-derivative does not.

   Subsequent work by Fry and his colleagues at Parke-Davis confirmed the usefulness of the tetrafluorotyrosine moiety as a nonphosphorylatable analog, in this case for peptides that target the epidermal growth factor receptor (EGFR) (Fry et al. 1994). The phenylalanine-containing "parent" peptide acetyl-Leu-Ala-Glu-Glu-Ser-Ala-Phe-Glu-Glu displays a K_i of 150 μM, whereas the corresponding l- and d-tetrafluorotyrosine-containing derivatives exhibit relative inhibitory enhancements of threefold and eightfold, respectively. The Parke-Davis group also prepared peptides that contained other l-tyrosine analogs, including 3-fluorotyrosine, 3-iodotyrosine, and d-tyrosine, but all of these derivatives were ineffective EGFR inhibitors. Curiously, 3-iodotyrosine was subsequently found to serve as an excellent tyrosine replacement in a cyclic peptide targeting Src (Alfaro-Lopez et al. 1998).

   Walsh, Cole, and their colleagues also examined the use of tetrafluorotyrosine as a tyrosine replacement in a C-terminal Src kinase (CSK)-targeted peptide (Cole et al. 1995; Kim and Cole 1998). However, in this case, the peptide serves as a substrate, rather than as an inhibitor, for CSK. These results suggest that the applicability of tetrafluorotyrosine as a nonphosphorylatable tyrosine replacement is kinase-dependent.

   Lam and his collaborators have prepared a series of active site-directed peptides that target the Src tyrosine protein kinase (Lou et al. 1997). These

investigators employed both d- and l-napthylalanine (Nal) derivatives in place of the phosphorylatable <u>tyrosine</u> moiety in the sequence Gly-Ile-<u>Tyr</u>-Trp-His-His-Tyr. The corresponding phenylalanine derivative was not prepared; however, the d-Tyr was, which gives a measure of the inherent affinity of the peptide for Src. The IC_{50} for Gly-Ile-d-<u>Tyr</u>-Trp-His-His is 50 µM, which indicates that the peptide framework is, comparatively speaking, a remarkably effective peptide-based inhibitor. The corresponding Gly-Ile-Nal-Trp-His-His derivative exhibits only a twofold improvement in IC_{50} relative the d-Tyr analog. However, the doubly substituted Gly-Ile-Nal-Trp-His-His-Nal exhibits an IC_{50} of 4 µM, suggesting that the C-terminal Nal is able to access sites outside of the immediate active site region. Interestingly, one of the less effective inhibitors Gly-Ile-Nal-Trp-His-His-Tyr (IC_{50}=27 µM) proved to be remarkably selective for Src versus other closely related members of the Src kinase family (Lyn and Lck; IC_{50}>1 mM).

One of the difficulties associated with the acquisition of nonphosphorylatable tyrosine surrogates is their synthesis, which typically resorts to the use of achiral starting material. Following a resolution step, the analogs must then be appropriately protected for use in solid phase peptide synthesis. Some of these difficulties have been circumvented by Kim and Cole, who employed the enzyme tyrosine phenol lyase to prepare gram quantities of an assortment of fluorinated tyrosine analogs (Kim and Cole 1998). The Lawrence group has developed a library-driven strategy, which allows one to prepare and subsequently screen a wide assortment of commercially available aryl-containing amines as peptide-based nonphosphorylatable tyrosine analogs (Niu and Lawrence 1997a,b). In spite of the fact that these are peptide derivatives, issues related to synthesis, resolution, and protection of these tyrosine substitutes are all bypassed.

Although the most common protein kinase peptide substrates possess a phosphorylatable residue embedded within the interior of the peptide, protein kinases will also phosphorylate peptides containing tyrosine, serine, and threonine moieties appended off the N- or C-terminus of these substrates. For example, Src catalyzes the phosphorylation of Arg-Arg-Arg-Arg-Arg-Leu-Glu-Glu-Leu-Leu-Tyr-amide (the arginine residues are present for assay purposes, not enzyme recognition). C- and N-terminal residues can be readily appended onto the active site-directed peptide after solid phase peptide synthesis. This allows one to employ potential tyrosine analogs that are not protected, possess functionality that might not survive the harsh conditions of peptide synthesis, and even lack the standard α-stereocenter. The synthetic strategy utilizes a solid phase peptide synthesis support (Kaiser's oxime resin) that allows the tyrosine analog to be attached to the synthesized peptide in a fashion that simultaneously promotes cleavage from the resin (Kaiser et al. 1989). For example, a wide assortment of phenylethylamine derivatives was attached to the C-terminus of a Src active site-directed peptide (Niu and Lawrence 1997a,b).

17   **18**

These were screened for inhibitory potency, and the lead analog was identified as the dopamine derivative **17** (which, in spite of a *p*-substituted aromatic alcohol, does not serve as a substrate). A peptide containing the amino acid analog of dopamine, l-dopa (**18**), was subsequently synthesized and shown to display an inhibitory potency (K=16 μM) that exceeds the parent phenylalanine-containing peptide by 60-fold.

Cole and his colleagues have reported a high-affinity bisubstrate analog for the insulin receptor protein kinase (IRK) (Parang et al. 2001).

Ac-Lys-Lys-Lys-Leu-Pro-Ala-Thr-Gly-Asp   Met-Asn-Met-Ser-Pro-Gly-Val-Asp

19

Compound **19** was designed based upon a dissociative mechanism for phosphoryl transfer in the IRK active site. The authors reasoned that the approximately 5 Å that separates the aromatic amine nitrogen from the γ-thiophosphate phosphorous roughly recapitulates the distance between acceptor and donor in a metaphosphate-like dissociative mechanism. Unlike the bisubstrate analogs reported to date for the serine/threonine-protein kinases, compound **19** serves as a competitive inhibitor versus both variable ATP and peptide substrate. The K_i of **19** is 370 nM, which corresponds to a binding energy that is roughly equal to the sum of the ATP- and protein-binding site portions of the inhibitor. As one might expect for an inhibitor that contains a peptide sequence targeting IRK, compound **19** is ineffective versus CSK (K_i~40 μM). These investigators also obtained the crystal structure of the inhibitor bound to the tyrosine kinase domain of IRK. The latter confirmed that the inhibitor is bound in a bisubstrate-like mode with the expected distance between the anilino nitrogen and the γ-phosphorous. In addition, the anilino nitrogen is engaged in a hydrogen bond to a key active site Asp residue.

Budde, McMurray, and their collaborators reported an N-myristoylated peptide, myr-Glu-Phe-Leu-Tyr-Gly-Val-Phe-Asp-amide, that serves as an apparent bisubstrate analog for Src (Ramdas et al. 1999). Surprisingly, the cor-

responding peptide with a free N-terminus is a Src substrate. However, upon acylation the substrate is converted into an inhibitor with the caveat that the acyl group must be lauryl ($C_{11}H_{23}CO-$) or longer. Clearly, the unexpected structure/activity relationship of this inhibitory species places it in an unusual category in that there is no obvious consolidated nonphosphorylatable tyrosine mimetic present in the peptide framework. Nonetheless, the fatty acyl-peptide serves as a competitive inhibitor with respect to both variable ATP and peptide substrate (poly Glu_4Tyr), thereby rendering it, like compound **19**, a bisubstrate analog. The nonacylated peptide itself blocks phosphorylation of poly Glu_4Tyr with a K_i of 260 μM via a competitive pattern. Myristic acid also serves as an inhibitor of the Src-catalyzed phosphorylation of poly Glu_4Tyr, but in this instance the fatty acid competes with ATP (K_i=35 μM).

The conjoined myr-Glu-Phe-Leu-Tyr-Gly-Val-Phe-Asp-amide exhibits bisubstrate inhibition with K_i values of 3 μM (variable Glu_4Tyr) and 6 μM (variable ATP). Consequently, the fatty acyl-peptide is unable to serve as a Src substrate because ATP is unable to bind to the active site in the presence of the myristyl group. Unfortunately, this inhibitory species does not display selectivity against other protein kinases (CSK, PKA, and the FGF receptor). However, it may ultimately be possible to enhance either selectivity or potency by placing the fatty acid moiety at different sites along the peptide chain to minimize the distance between the site of phosphorylation (i.e., the Tyr residue) and the ATP-binding site. These results stand in interesting contrast to earlier work described by Ward and O'Brian (O'Brian et al. 1990; O'Brian et al. 1991; Ward and O'Brian 1993). PKC peptide substrates, upon *N*-myristoylation, are converted into inhibitors with IC_{50} values of between 3–10 μM, depending upon the amino acid sequence. However, unlike the myristoylated peptides that inhibit Src, Ward and O'Brian's PKC inhibitors do not display a competitive pattern versus variable ATP, nor do they serve as competitive inhibitors versus peptide substrate. The authors conclude that their myristoylated peptides bind to a different enzyme form (i.e., the free enzyme) than the peptide substrate (i.e., the enzyme-ATP complex). This accounts for the noncompetitive inhibition pattern versus peptide substrate. In addition, Ward, O'Brian, and their colleagues have suggested that the inhibitory effect conferred by the myristoyl appendage is due, at least in part, to its interaction with the phosphatidylserine cofactor of PKC (O'Brian et al. 1990). However, more recent studies suggest that a myristyl-binding region is located in close proximity to the peptide-binding region of the active site (Zaliani et al. 1998). Consequently, it appears likely that the inhibitory behavior of myristoylated peptides toward PKC is at least partly due to the presence of a near active site hydrophobic region that is able to accommodate the lipophilic fatty acid moiety.

4.1.3
Phosphotyrosine Analogs

SH2 and PTB domains, as well as protein tyrosine phosphatases, recognize sequences that encompass phosphotyrosine. However, since the latter is hydrolyzable, peptides that contain this residue serve only as transient antagonists of protein–protein interactions. A number of nonhydrolyzable phosphotyrosine mimetics have been described, in large part due to the effort of Burke and his collaborators at the NIH. Since phosphotyrosine analogs have been the topic of recent reviews, only a cursory discussion of their current status will be presented (Burke et al. 2001; Burke and Lee 2003).

20   **21**   **22**
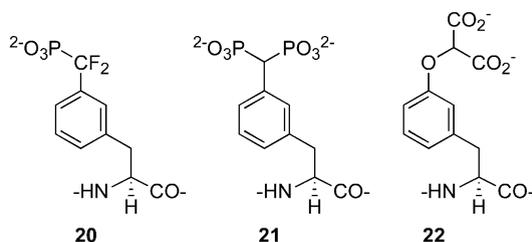
Perhaps the most extensively employed of all phosphotyrosine analogs is the difluorophosphonate derivative **20**. The synthesis of this analog was first described by Burke and his colleagues (Smyth et al. 1992; Burke et al. 1993) and was subsequently introduced into peptides and evaluated as SH2 ligands in 1994 (Burke et al. 1994; Gilmer et al. 1994). The phosphonate moiety is not only resistant to hydrolysis by adventitious intracellular phosphatases, but the electron-withdrawing effect of the difluoro substituents ensures that the phosphonate is doubly ionized at physiological pH. In addition to the difluoromethylene moiety, a variety of other groups have been used to link the phosphoryl moiety to the aromatic nucleus, including methylene, hydroxymethylene, and fluoromethylene (Burke et al. 1994). The $-OPO_3H$ group on phosphotyrosine has also been replaced with $-CH_2COOH$ (Gilmer et al. 1994; Tong et al. 1998), $-OCH_2COOH$ (Burke et al. 1999), $-CHOHCOOH$ (Beaulieu et al. 1999), $-CF_2COOH$ (Burke et al. 1999), $-CH_2SO_3H$ (Gilmer et al. 1994), $-CH_2CH_2COCOCH_3$ (Gilmer et al. 1994), and NO_2 (Gilmer et al. 1994). In addition, the phosphoryl group has been directly attached to the aromatic ring (i.e., no bridging group) (Stankovic et al. 1997). A number of geminal analogs [e.g., **21** (Bohacek et al. 2001) and **22** (Kole et al. 1995; Ye et al. 1995)] have been prepared and examined as well. Although the general statement that peptides containing **20** in place of phosphotyrosine possess the greatest affinity for their intended targets may be true, there is sufficient variability among the protein-binding domains that any generalization concerning binding preferences is dangerous. For example, the monocarboxylic acid analogs of phosphotyrosine appear to have reasonable affinity for cer-

tain SH2 domains, but are poor inhibitors for the protein tyrosine phosphatase family member PTP1B (Gao et al. 2000).

4.1.4
Proline Analogs

Proline is commonly thought of as a turn-promoting residue. Indeed, a large number of turn-inducing proline mimetics have been described, some of which have been synthetically incorporated into peptides. Sarcosine, also known as N-methyl glycine (CH_3NHCH_2COOH), is a commonly employed proline replacement. Indeed, the latter has been used in substrates for cyclin-dependent protein kinases, enzymes that require a proline residue on its intended substrate to direct phosphorylation to an adjacent serine moiety. For example, Ando and coworkers showed that peptides containing the sequence from vimentin, Leu-Gly-Ser-Ala-Leu-Arg-Arg-Arg-amide, in which the alanine moiety is replaced by either proline or sarcosine, serve as enhanced substrates for the cdc2 protein kinase (V/K 600-fold for Pro-peptide and 90-fold for Sar-peptide versus the Ala-peptide) (Ando et al. 1993). These investigators subsequently examined the specificity of the cdc2 and cdk5 proline-directed protein kinases with respect to a variety of N-substituted proline replacements including, N-methyl alanine, N-methyl leucine, N-methyl valine, and the four (azetidine carboxylic acid) and six ring (piperidine carboxylic acid) analogs of proline (Ando et al. 1997). Perhaps not too surprisingly, the peptides containing the ring analogs of proline were among the best substrates for these protein kinases, with the proline-containing parent serving as the most efficient substrate. Nevertheless, the sarcosine derivative serves as a substrate as well, albeit somewhat more efficiently for the cdc2 versus cdk5 enzyme. The latter observation is significant since sarcosine is the parent of N-substituted glycine derivatives. Peptides containing the latter ("peptoids") are readily prepared using standard solid-phase peptide synthesis conditions (Figliozzi et al. 1996; Burkoth et al. 2003). Indeed, recent improvements have reduced the coupling time to a few minutes (Olivos et al. 2002). Consequently, it is now possible to prepare large libraries of peptoid-based derivatives in an essentially automated fashion.

Lim and his colleagues have reported that N-substituted amino acids serve as proline replacements in peptide-based ligands targeting both SH3 and WW domains (Nguyen et al. 1998; Nguyen et al. 2000). These investigators demonstrated that the proline selectivity for these domains is due to a preference for N-substituted residues and not simply a function of the rigid cyclic structure that is unique to proline. The SH3 domain from Sem5 recognizes the motif PPPVPPR, whereas the WW domain from Yap prefers GTPPPPYTVG (where the "essential" prolines are underlined). The initial strategy employed a scanning approach in which both essential and nonessential prolines were replaced by either the C^α-substituted alanine or

the N-substituted sarcosine. Replacement of the essential prolines with sarcosine is well tolerated, whereas introduction of alanine at these sites is not. The three-dimensional structures of these protein recognition domains are known. The ligand binds as a polyproline type II left-handed helix, which contains three residues per turn. The binding grooves on these protein domains can accommodate substituents from the type II helix, but only in a closely packed fashion. The latter can be achieved by a motif in which a C^{α}-substituent is adjacent to an N-substituted residue. A small library was prepared that consisted of peptides containing 11 different N-substituted glycines at the two essential proline sites in an SH3-targeted ligand. These were screened against the SH3 domains from Src, Grb2, Crk, and Sem5.

TyrGluValProProPro—N—CO-ProArgArgArg

23

TyrGluValProProPro—N—CO-ProArgArgArg OCH_3 CH_3O

24
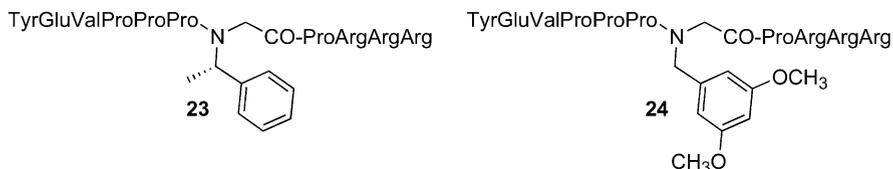
The lead peptide for Grb2, **23**, exhibits a K_D of 40 nM, more than 100-fold better than that of the corresponding proline-containing derivative. A subsequently prepared library, containing N-substituted glycine residues with oxygen- and nitrogen-containing substituents, was also evaluated (Nguyen et al. 2000). The latter functionality was chosen for their ability to interact with polar residues that reside near the ligand-binding grooves of the SH3 domain. Compound **24** exhibits a 30 nM K_D for the Grb2 SH3 domain and high selectivity versus Crk (300-fold) and Src (1,500-fold). A high-affinity ligand with impressive selectivity was also identified for the Crk SH3 domain. Double proline-substituted derivatives were also prepared, but these ligands failed to show an additive improvement in affinity for their protein targets. The latter is likely because the N-substituted glycines at the two different "essential" proline sites were identified independently of one another. Introduction of an N-derivatized glycine at one position likely induces a conformational change between ligand and protein that influences the interaction at the second proline site. Nevertheless, these results demonstrate that "essential" residues cannot only be replaced with unnatural analogs, but that the latter can furnish dramatically enhanced potency and selectivity.

4.2
Multidomain-Targeting Peptides

Many protein kinases contain two or more independent protein-binding domains. This includes the best-known example, Src, which contains SH1 (active site), SH2, and SH3 domains. Others, such as the LIM-kinase (LIM, PDZ, and active site regions) likewise contain multiple domains (Stanyon

and Bernard 1999). The twin issues of potency and specificity can potentially be addressed by simultaneously targeting two protein interaction domains on the same enzyme. First, a single bivalent peptide that concomitantly interacts with two binding domains should display an enhanced affinity for the protein kinase target relative to either monovalent peptide alone. Second, although the sequence homology within closely related domains can be very high, the relative three-dimensional disposition of two protein interaction domains in the intact protein may differ substantially from one protein kinase to the next. Consequently, the linker that connects the two monovalent consensus sequences could play an important role in conveying enzyme selectivity. At this point in time, however, the potential of bivalent (or multivalent) peptides that can interact with two or more protein interaction domains is at an early stage in development and has not yet been fully realized.

One of the first papers to describe the preparation of bivalent ligands for signaling proteins dealt with the SH-PTP2 protein tyrosine phosphatase (Pluskey et al. 1995). The latter contains two SH2 domains. Occupancy of these domains is known to stimulate phosphatase activity. Previous studies had identified peptides SLNpYIDLDLVK and LSTpYASINFQK that specifically target the N-terminal and C-terminal SH2 domains, respectively (Case et al. 1994; Sugimoto et al. 1994). These monovalent species were linked with one another via an aminocaproic acid bridge to create a consolidated derivative, LNpYIDLDLVK-(6-aminocaproic acid)$_4$-LSTpYASINFQK, that can simultaneously interact with both SH2 domains. The latter does not act as an inhibitor, since it was designed to promote phosphatase activity. Nevertheless, it stimulates the phosphatase by 37-fold, compared with the 9- to 16-fold displayed by the monovalent consensus peptides. In addition, the heterodimeric peptide displays a 60- to 90-fold higher affinity for SH-PTP2 than either monomer peptide alone.

Cowburn, Barany, and their colleagues have described peptides that simultaneously target the SH2 and SH3 domains of Abl (Cowburn et al. 1995). The SH2-targeting sequence (PVpYENV-amide; K_D=2.0±0.2 μM) was attached to the SH3-targeting sequence (PPAYPPPPVP; K_D=10.5±0.2 μM) via a lysine side chain as illustrated in **25**. The latter exhibits a K_D of 249±5 nM for the SH2–SH3 domain construct, which is approximately 10- to 80-fold greater than that of individual monovalent consensus peptides. These investigators subsequently prepared a series of bivalent ligands containing the individual monomers in the four possible orientations illustrated in Scheme 1 25*A*–*D* (Xu et al. 1999). The highest affinity ligand, **25***C*, exhibits a K_D of 190 nM for the Abl SH2-SH3 (where LINKER=Gly$_7$K).

SH2-directed

PVpYENV-GGGGGG-HN

H_2N

SH3-directed   *SH3-directed*   *SH2-directed*

PPAYPPPPVP-HN   amide   PPAYPPPPVP-GGGGGGG-HN   NH-PVpYENV-amide

H   H

O   O

25A   **25C**

N-terminus	C-terminus			C-terminus	N-terminus
SH2-peptide		LINKER		SH3 peptide	
		A			

N-terminus	C-terminus			N-terminus	C-terminus
SH3-peptide		LINKER		SH2 peptide	
		C			

C-terminus	N-terminus			N-terminus	C-terminus
SH2-peptide		LINKER		SH3 peptide	
		B			

N-terminus	C-terminus			N-terminus	C-terminus
SH2-peptide		LINKER		SH3 peptide	
		D			

Roques, Garbay, and coworkers prepared peptides that interact with the two SH3 domains of Grb2 (Cussac et al. 1999). As with the bivalent analogs described in this section, the strategy employed coupling two monomer units via a linker. Two SH3-binding monomer units, Val-Pro-Pro-Val-Pro-Pro-Arg-Arg-Arg, were attached to each other via their C-termini using the lysine-based strategy described by Cowburn and Barany. The bivalent species exhibits an affinity (K_D=40 nM) that is two to three orders of magnitude greater than that exhibited by the corresponding monomer (N-terminal SH3 domain: K_D=2.6 μM; C-terminal SH3 domain: K_D=40 μM). These investigators demonstrated that the peptide dimer blocks Grb2-Sos complex formation in cell lysates, can selectively pull down Grb2 from lysate versus other double SH3 domain-containing proteins, and interferes with neurite formation in nerve growth factor-treated PC12 cells.

Miller and his coworkers prepared a series of SH2 domain-assisted *substrates* for the Abl tyrosine kinase (Pellicena et al. 1998). These investigators found that the presence of an SH2 domain binding sequence appended to the active site-directed sequence enhances substrate efficacy, specifically via a tenfold reduction in K_m. For example, a peptide of the general structure (active site-directed peptide)–(SH2-directed peptide) displays a K_m of 69±11 μM and a V_{max} of 3.0±0.1 μmol/min-mg. The key residue in the SH2-directed component is phosphoTyr, which is essential for SH2 recognition. When a Phe replaces this critical residue, the peptide displays a significantly larger K_m of 680±90 μM, yet a V_{max} (3.0±0.1 μmol/min-mg) that is essentially unaltered. A peptide containing an inversed orientation, namely (SH2-directed peptide)–(active site-directed peptide), likewise displays a relatively low K_m value (72±5 μM), which is presumably a reflection of enhanced affinity via coordination to the SH2 domain.

Lawrence and his collaborators employed the multidomain targeting approach to create combined active site/SH2 domain-directed inhibitors of Src (Profit et al. 1999; Profit et al. 2001). Peptide-based inhibitors of tyrosine kinases must confront at least two challenges. First, peptides that target the ac-

tive site tend to be exceedingly poor inhibitory agents. Second, although peptides that bind to the SH2 domain are of reasonably high affinity (~low µM) they have the unintended consequence of activating the kinase. Indeed, by analogy, Shoelson's and Walsh's bivalent ligand, which binds to the two SH2 domains of SH-PTP2, dramatically enhances phosphatase activity. However, an inhibitor that combines SH2 and active site-binding properties should simultaneously display enhanced affinity and inhibitory potency. In addition, bivalent analogs can furnish an assessment of the distance and spatial relationship between the protein interaction domains on the protein under evaluation. Indeed, the Lawrence team referred to their small library (16 compounds) as "molecular rulers." In a manner analogous to Barany and Cowburn, a series of differentially oriented SH2-directed and active site-directed bivalent ligands were prepared. γ-Aminobutyric acid (GABA) residues were employed in the linker region to connect the monovalent ligands. The active site-directed fragment, -Glu-Glu-Leu-Leu-(F_5Phe)-, contains pentafluorophenylalanine, which had been previously identified as a nonphosphorylatable tyrosine surrogate (*vide supra*) (Niu and Lawrence 1997a). The SH2 sequence, -pTyr-Glu-Glu-Ile-, was based on the well-known preference of the SH2 domain from Src for a sequence motif containing pTyr followed by at hydrophobic amino acid at the P+3 position (Songyang et al. 1993).

SH2-directed   *A.S.-directed*

acetyl-**pTyr**-Glu-Glu-Ile-Glu-(GABA)$_n$-Glu-Glu-Leu-Leu-**F_5Phe**-amide

26A

A.S.-directed   *SH2-directed*

acetyl-Glu-Glu-Leu-Leu-**F_5Phe**-(GABA)$_n$-**pTyr**-Glu-Glu-Ile-Glu-amide

26B

N-terminus C-terminus		N-terminus C-terminus		C-terminus N-terminus		N-terminus C-terminus
SH2-peptide	GABA$_n$	A.S. peptide		SH2-peptide	Ga-GABA$_n$	A.S. peptide
A				**C**		

N-terminus C-terminus		N-terminus C-terminus		N-terminus C-terminus		C-terminus N-terminus
A.S.-peptide	GABA$_n$	SH2 peptide		A.S.-peptide	LINKER	SH2 peptide
B				**D**		

Of the four possible relative orientations between the SH2- and active site-directed fragments (Scheme 2), only three were prepared (*A–C*) since the fourth (*D*) was subsequently ruled out as suboptimal (*vide infra*). The orientation illustrated in *C* requires that the peptide chains in the two monovalent units run anti-parallel to one another. This reversal of chain polarity was achieved via insertion of a glutaric anhydride (Ga) residue between the SH2-directed sequence and the GABA linker. Orientation *B* furnished the most effective inhibitors possessing the shortest linkers (optimal: GABA$_n$= GABA$_3$). In this particular case, the tyrosine surrogate F_5Phe (occupying the

active site) lies only three residues from the pTyr moiety which resides in the SH2 domain (see **26***B*).

This suggests that the active site region and the SH2 domain are situated close to one another in the active form of the enzyme. The series **26***D* was not prepared since, even in the absence of a GABA linker, the key F_5Phe and pTyr moieties are positioned at least five residues apart. The latter is suboptimal relative to the relationship in **26***B*. The lead bivalent analog **26***B* (where $GABA_n$=$GABA_3$) displays an enhanced 120-fold inhibitory potency relative to the simple active site-directed monomer (IC_{50}=13±1 μM versus IC_{50}= 1590±170 μM, respectively).

4.3
Structural Modification of Consensus Sequence Peptides

4.3.1
Conformationally Biased Peptides

Cyclization of peptides improves stability against proteases while affording conformational constraints that may enhance inhibitory potency. Research teams at M.D. Anderson and the University of Arizona have explored the efficacy of this structurally distinct class of active site-directed protein kinase inhibitors.

Lam and his colleagues have shown that the Tyr-Ile-Tyr-Gly-Ser-Phe-Lys-amide motif serves as an effective Src substrate (Lam et al. 1995) and that the peptide might bind to the active site in the form of a *β*-turn (Lou et al. 1997). In order to assess the reasonableness of this notion, Lam, Hruby, and their coworkers prepared a series of disulfide bridge-cyclized inhibitory peptides that target Src (Alfaro-Lopez et al. 1998). In addition, these investigators took advantage of the observation that 2-naphthylalanine (Sect. 11) serves as a nonphosphorylatable tyrosine mimetic (Wu et al. 1996).


27   28
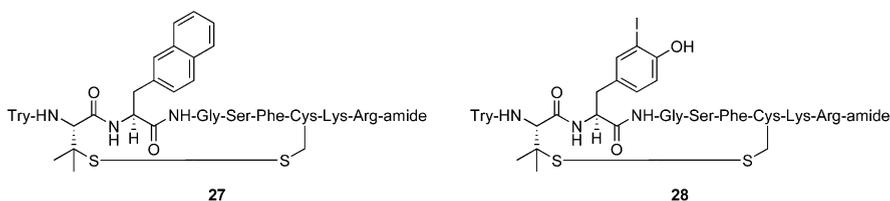
Compound **27** exhibits an IC_{50} of 1.6 μM. Interestingly, and somewhat surprisingly, potency was further improved by tenfold when the naphthylalanine was replaced with 3-iodotyrosine (**28**). This observation stands in marked contrast to the ineffectiveness of 3-iodotyrosine as a tyrosine mimetic in an EGFR-targeted inhibitory peptide (Fry et al. 1994). Peptide **28** is not only an effective Src kinase inhibitor but also exhibits impressive selectivity in favor of Src versus other Src kinase family members (20-fold versus

Lck and >1,300-fold versus Lyn). In addition, these investigators found that several of their less potent Src kinase inhibitors (IC_{50} values 1–3 µM range) exhibit even better selectivity profiles than **28**.

McMurray, Budde, and coworkers likewise examined the effectiveness of cyclic peptide inhibitors on the Src kinase. These investigators first examined the affinity of a series of cyclic peptide substrates for Src (i.e., the ability of these substrates to block the phosphorylation of poly Glu_4Tyr) (McMurray et al. 1998). One of the lead compounds, cyclo[Asp-Asn-Glu-Tyr-Ala-Phe-Phe[7]-Gln-D-Phe-Pro) displays a K_i of 150 µM, which is nearly identical to its Michaelis constant as a substrate (140 µM). Insertion of an arginine residue at position 7 resulted in a dramatic loss in enzyme affinity, whereas a glutamic acid residue at this site is well tolerated. On the basis of this observation, these investigators concluded that residues at this site are positioned within a positively charged region of the enzyme. Indeed, when Phe[7] was subsequently replaced with a series of 14 different analogs, the lead inhibitors contained a negatively charged residue at this position [4-carboxyphenylalanine (K_i=0.85 µM) and phosphotyrosine (K_i=1.1 µM)] (Wang et al. 2000). The carboxylphenylalanine-containing cyclic peptide displays an impressive selectivity profile in favor of Src (>100-fold against Yes; >300-fold versus Lck; >1,000-fold versus PKA; 1,200-fold versus FGF receptor; 1,800-fold versus Abl; >2,000-fold versus CSK).

Watterson and his group described a different strategy for topologically biasing an active site-directed peptide (Lukas et al. 1999). Myosin light chain kinase (MLCK) was the target in this particular case. These investigators first identified a nonapeptide sequence, Arg-Lys-Lys-Tyr-Lys-Tyr-Arg-Arg-Lys-amide that exhibits both a remarkable affinity (IC_{50}=50 nM) and selectivity (~4,000-fold versus CaM kinase II) for MLCK. Based on the screening results of closely related peptides, in combination with molecular modeling, it was proposed that the peptide might bind to the active site region in an extended conformation. The structural constraints inherent within 4-aminocyclohexanecarboxylic acid were used to promote this desired conformation by inserting the residue at specific sites within the peptide sequence.


Arg-Lys-Lys-Tyr—N(H)— ... —C(=O)—Tyr-Arg-Arg-Lys-amide

29
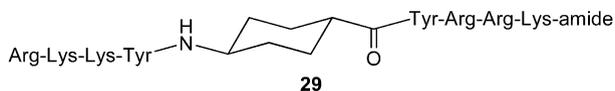
The most potent of these derivatives (**29**) exhibits an IC_{50} (40 nM) similar to that of the parent nonapeptide. Although improved potency and selectivity were not observed versus the already formidable peptide parent, the results in this study suggest that conformational constraints could serve as scaffolds upon which an array of functionality can be appended.

4.3.2
Terminally Modified Peptides

The fatty acid-modified peptides described in Sect. 4.1.2 are examples of terminally modified consensus sequences. Addition of the long alkyl chain furnishes enhanced active site affinity via coordination to ancillary binding pockets. However, it should be possible to access potential sites of interaction that are extensions of the protein-binding pocket with non-natural substituents. Schreiber's group explored this concept utilizing N-terminally modified peptides that target the SH3 domain of Src (Combs et al. 1996). The three-dimensional structure of the protein was employed as a guide to focus structural diversity into a potential binding region adjacent to the site at which a peptide ligand of the SH3 domain is known to reside. Structural diversity was created using a split-and-pool approach. Pro-Leu-Pro-Pro-Leu-Pro-resin was split into 33 equal amounts and each fraction subsequently modified at the peptide N-terminus with one of 32 different non-natural amino acids (plus no residue at all). The fractions were recombined and the process repeated two additional times. The one-bead/one-compound library was then exposed to a phosphatase-modified SH3 domain. Beads possessing high-affinity ligands were visually identified under the microscope using the phosphatase as a colorimetric reporter.
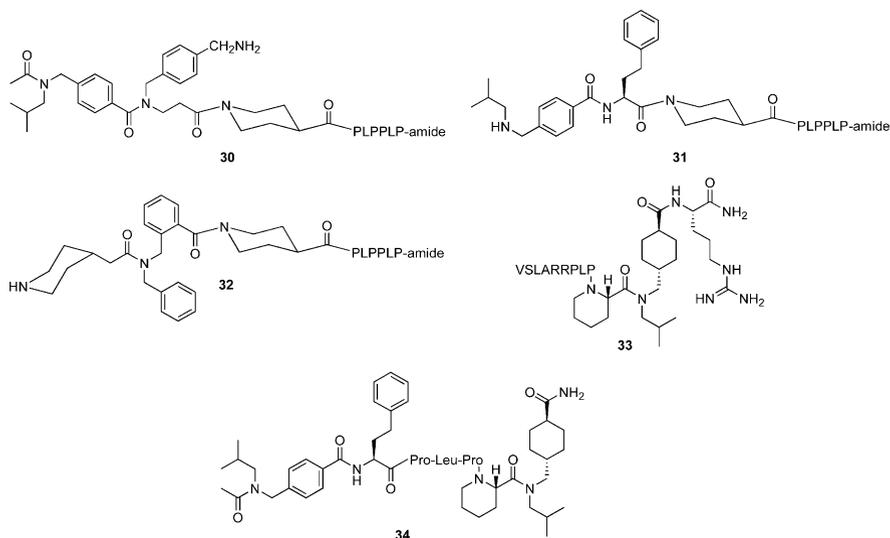
The lead peptide identified in this study, **30**, exhibits a K_D of 3.4 µM for the Src SH3 domain and is approximately 50-fold more selective for Src versus the corresponding SH3 domain from PI3 kinase. By comparison, the parent peptide, acetyl-Pro-Leu-Pro-Pro-Leu-Pro, exhibits a K_D of greater than 1 mM for Src's SH3 domain. A subsequent study, using a slightly larger set

of non-natural monomers and greater structural diversity at each position, identified additional ligands for the Src SH3 domain (**31**; K_D=0.9 μM) as well as leads for the closely related SH3 domain of Hck (**32**; K_D=1.0 μM) (Kapoor et al. 1998). Some of these ligands proved to be moderately selective (e.g., **32** is 46-fold more selective for Hck than Src).

The Schreiber group extended the concept of structural diversity to the C-terminus of the SH3-directed ligand as well (Morken et al. 1998). In this instance, the library took the form of Val-Ser-Leu-Ala-Arg-Arg-Pro-Leu-Pro-M3-M2-M1-resin, where M3, M2, and M1 represent 50 different residues, which included 49 monomers plus an omitted residue. The goal of this work was to identify a replacement for the C-terminal Leu-Pro dyad in the parent peptide -Pro-Leu-Pro-Pro-Leu-Pro-. The lead **33** displays a K_D of 2.6 μM for the SH3 domain from Src. A portion of the C-terminal non-natural component from **33** was then appended, along with the N-terminal nonnatural component in **31**, to the Pro-Leu-Pro core (**34**). The latter compound displays a K_D of 1.1 μM for Src's SH3 domain.

4.3.3
Globally Modified Peptides

Protein interaction domains have evolved to accommodate specific sequences of preferred amino residues on their protein-binding partner. However, each amino acid site on the bound sequence is limited to a genetically encoded molecular diversity of 20. In reality, the latter is slightly larger due to a small array of possible post-translational modifications. Nevertheless, it is not difficult to imagine that there exist a wealth of potential binding interactions that lie just outside of the reach of this limited set of naturally occurring residues. One strategy to enhance molecular diversity is to prepare a wide assortment of Fmoc and side chain-protected, unnatural, amino acid derivatives and then synthesize the corresponding library of peptides. However, a reasonable 50-fold enhancement in molecular diversity vis-à-vis genetically encoded residues would require the preparation of 1,000 different monomers, a nontrivial task.

Lawrence and his colleagues outlined an approach that creates high diversity at any desired site along a peptide chain using readily available carboxylic acid-containing compounds (Lee and Lawrence 1999). The strategy, as outlined in Scheme 3, employs a consensus sequence peptide containing 2,3-diaminopropionic acid ("Dap"), appropriately inserted at key sites along the peptide chain.
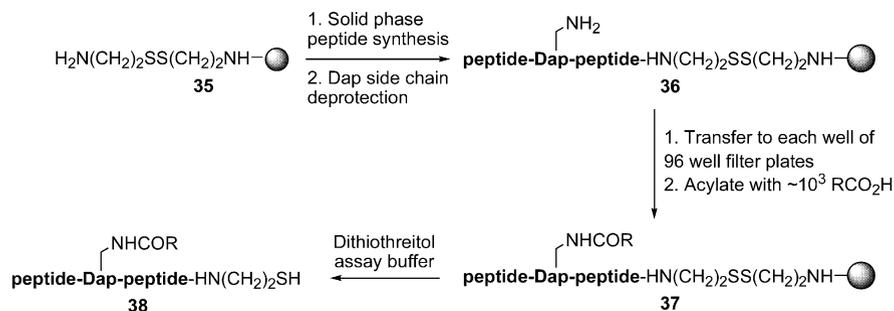
The peptide is synthesized on a disulfide-substituted Tentagel resin. Once prepared, the peptide-resin is distributed in equal amounts to the individual wells of a 96-multiwell plate designed for organic synthesis (i.e., the bottom of each well contains a frit that allows multiple washings without loss of the peptide-resin). Each well is then exposed to one of approximately 1,000 different carboxylic acid compounds. In short, the library is prepared in parallel, thereby obviating the necessity of molecular deconvolution. Once the modification in the Dap residue is complete, any side chain protecting groups on the peptide are removed with trifluoroacetic acid. After multiple washings to remove residual acid, the individual peptides are cleaved from the resin with assay buffer, which contains dithiothreitol. The peptides can then be directly assessed for potency.

Lawrence and his team employed a structure-based strategy, in combination with the synthetic approach outlined in Scheme 3, to identify high-affinity ligands for the SH2 domain from Lck (Lee and Lawrence 1999, 2000; Yeh et al. 2001). The three-dimensional structure of the Lck SH2 ligand, acetyl-pTyr-Glu-Glu-Ile-amide, bound to its protein target, had been previously reported (Tong et al. 1996). Three sites on the ligand, the N-terminal acetyl moiety and the two Glu side chains, are oriented into regions of the SH2 surface that could potentially accommodate modified analogs of the acetyl and glutamic acid moieties.
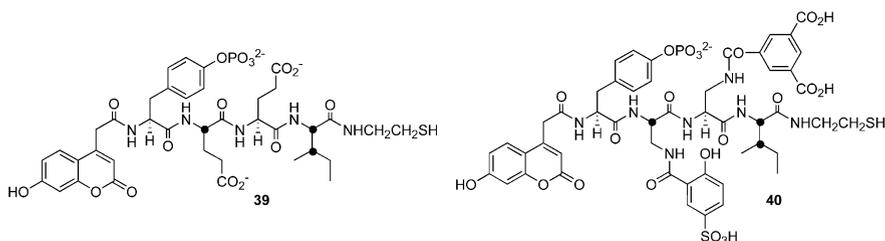
The initial library and its subsequent screen furnished compound **39**, which contains a coumarin moiety in place of the former N-acetyl group. The coumarin-derivatized peptide exhibits a K_D of 35 nM for the Lck SH2 domain, approximately two orders of magnitude greater than the parent

peptide aceyl-pTyr-Glu-Glu-Ile-amide. Subsequent identification of the glutamic acid replacements furnished **40**, which displays a K_D of 200 pM, approximately four orders of magnitude better than the starting consensus peptide. An analogous approach was recently used to construct an inhibitor for the α isoform of PKC (Lee et al. 2004). The inhibitor displays a K_i of 800 pM and a selectivity of greater than 400-fold versus other conventional, novel, and atypical PKC isoforms.

Lawrence, Zhang, and their colleagues reported a variation on the Scheme 3 strategy that provided a high-affinity inhibitor for PTP1B (Shen et al. 2001). Once again, a structure-based approach was employed that directed molecular diversity toward potential binding sites on the target protein surface. In this particular case, PTP1B had been previously shown to bind phosphotyrosine at two distinct sites, one at the active site and the other at a position adjacent to the active site (Puius et al. 1997).

peripheral site-targeted   **41**   active site-targeted   **42**
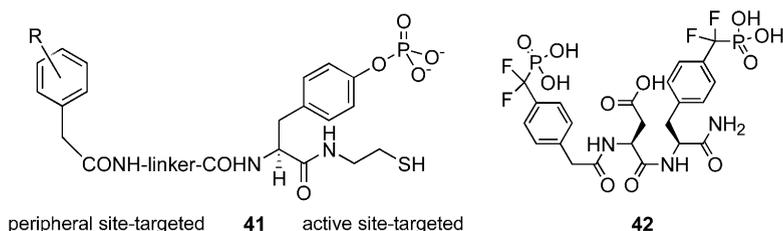
A library of the general form **41** was prepared using the disulfide Tentagel resin **35**. Molecular diversity was inserted at the N-terminal and linker positions. The lead compound **42** exhibits a K_i of 2.6 nM for PTP1B and a selectivity of between 1,000- and 10,000-fold versus a panel of fifteen other protein phosphatases.

5
Summary

An exceedingly important, but time-consuming area of drug design is the conversion of consensus recognition sequences into small molecules with drug-like properties. The acquisition of peptidomimetics requires a combination of detailed structural information of the target protein, an intensive synthetic effort, and gifted insight. The rapid development of human immunodeficiency virus (HIV) protease inhibitors stands as a testimony to the fact that, given enough resources, it is possible to successfully create potent small molecule inhibitors that target protein–protein interaction sites (Abdel-Rahman et al. 2002). However, the sheer number of protein kinases rules out the kind of large-scale assault that transpired in the HIV arena on the protein kinase family as a whole. Nevertheless, at a minimum, a worthwhile

goal is the acquisition of high-affinity reagents for as many signaling proteins as possible. Although consensus sequences represent an obvious starting point, their transmogrification into high-affinity ligands remains an ongoing struggle. Given the large number of possible targets, simple rules or strategies for the conversion of modest-binding peptides into high-affinity reagents are a much sought after commodity. Amino acid analogs, which serve as high-affinity replacements for their natural counterparts, represent one such approach (Sect. 4.1). Alternatively, the notion of targeting two or more protein–protein interaction domains on a single protein kinase, represents a decidedly different tactic (Sect. 4.2). Finally, structurally modified consensus sequences that are either topologically biased or are able to access sites simply unavailable to standard amino acid residues represents a third strategy (Sect. 4.3). Inherent within all of these approaches is the possibility of general applicability to the family of protein kinases. Unfortunately, the route from peptide to a high-affinity species with equally high selectivity is often anything but straightforward. The disadvantage with protein kinases is their large number. However, this disadvantage is also an advantage in that they are all closely related. Consequently, inhibitor design rules that emerge from the study of a few representative members of this large enzyme family may ultimately prove applicable to the family as a whole.

References

Abdel-Rahman HM, Al-karamany GS, et al (2002) HIV protease inhibitors: peptidomimetic drugs and future perspectives. Curr Med Chem 9:1905–22

Adams JA, Taylor SS (1992) Energetic limits of phosphotransfer in the catalytic subunit of cAMP-dependent protein kinase as measured by viscosity experiments. Biochemistry 31:8516–22

Alfaro-Lopez J, Yuan W, et al (1998) Discovery of a novel series of potent and selective substrate-based inhibitors of p60c-src protein tyrosine kinase: conformational and topographical constraints in peptide design. J Med Chem 41:2252–60

Ando S, Tsujimura K, et al (1993) Phosphorylation of synthetic vimentin peptides by cdc2 kinase. Biochem Biophys Res Commun 195:837–43

Ando S, Ikuhara T, et al (1997) Role of the pyrrolidine ring of proline in determining the substrate specificity of cdc2 kinase or cdk5. J Biochem [Tokyo] 122:409–14

Beaulieu PL, Cameron DR, et al (1999) Ligands for the tyrosine kinase p56lck SH2 domain: discovery of potent dipeptide derivatives with monocharged, nonhydrolyzable phosphate replacements. J Med Chem 42:1757–66

Bohacek RS, Dalgarno DC, et al (2001) X-Ray structure of citrate bound to Src SH2 leads to a high-affinity, bone-targeted Src SH2 inhibitor. J Med Chem 44:660–3

Burke TR Jr, Smyth MS, et al (1993) Preparation of fluoro-4-(phosphonomethyl)-D,L-phenylalanine and hydroxy-4-(phosphonomethyl)-D,L-phenylalanine suitably protected for solid phase synthesis of peptides containing hydrolytically stable analogues of O-phosphotyrosine. J Org Chem 58:1336–40