Mutant Tyrosine Kinases with Unnatural Nucleotide Specificity Retain the Structure and Phospho-Acceptor Specificity of the Wild-Type Enzyme

Laurie A. Witucki,[1,5] Xin Huang,[2] Kavita Shah,[1,6]
Yi Liu,[1] Saw Kyin,[3] Michael J. Eck,[2]
and Kevan M. Shokat[1,4]
[1] Department of Cellular and
    Molecular Pharmacology
Box 0450
University of California, San Francisco
San Francisco, CA 94143-0450
[2] Department of Biological Chemistry and
    Molecular Pharmacology
Harvard Medical School, and
Department of Cancer Biology
Dana Farber Cancer Institute
44 Binney Street
Boston, MA 02115
[3] Department of Molecular Biology
Synthesis and Sequencing Facility
Princeton University
Princeton, NJ 08544
[4] Department of Chemistry
University of California, Berkeley
Berkeley, CA 94720

Summary

The direct substrates of one protein kinase in a cell can be identified by mutation of the ATP binding pocket to allow an unnatural ATP analog to be accepted exclusively by the engineered kinase. Here, we present structural and functional assessment of peptide specificity of mutant protein kinases with unnatural ATP analogs. The crystal structure (2.8 Å resolution) of c-Src (T338G) with N^6-(benzyl) ADP bound shows that the creation of a unique nucleotide binding pocket does not alter the phospho-acceptor binding site of the kinase. A panel of optimal peptide substrates of defined sequence, as well as a degenerate peptide library, was utilized to assess the phospho-acceptor specificity of the engineered "traceable" kinases. The specificity profiles for the mutant kinases were found to be identical to those of their wild-type counterparts.

Introduction

The Src family of protein tyrosine kinases catalyzes the phosphorylation of protein tyrosine residues and thereby plays an important role as a signaling switch in a variety of cellular pathways, including regulation of the cell cycle, antigen-dependent immune response, ion-channel regulation in the brain, vesicle uptake, and many others [1–6]. Existing genetic and biochemical methods are not able to unambiguously define the direct cellular sub-

strates of protein tyrosine kinases. This is due to the highly interconnected nature of these signaling pathways as well as the overlapping substrate specificity and redundancy of the protein kinase family [2, 7]. To deconvolute protein kinase signaling cascades, we have developed a chemical tagging approach that allows for identification of the direct substrates of protein kinases ([8–13]; Shaw & Shokat, this issue of *Chemistry & Biology*, pp. 35–47). This method requires a mutant kinase that has an altered active site into which a chemically modified ATP analog, A*TP (Figure 1), can bind and serve as a unique substrate for the mutant enzyme. The A*TP analog is chosen so as not to be a substrate for *any* wild-type protein kinase, a feature termed "orthogonality." Thus, only the direct substrates of the analog-specific (termed "as") kinase become radiolabeled. This strategy allows for exclusive tagging of the mutant kinase's direct substrates when [γ-^{32}P] A*TP is used as the labeling species in a cell or cell lysate.

Tyrosine kinases are bisubstrate enzymes containing a nucleotide (ATP) binding pocket and a protein or phospho-acceptor binding region. The question addressed in this study is whether alteration of the structure of one substrate (ATP → A*TP) in tandem with a complimentary mutation in the enzyme leads to alteration of the mutant enzyme's peptide (phospho-acceptor) specificity. Since the "traceable" mutant kinase is to be used to identify protein substrates that cannot be unambiguously confirmed by any other technique, it is imperative to critically assess the similarities/differences between the phospho-acceptor specificity of the wild-type and analog-sensitive kinases.

We present structural analysis of the cocrystal of c-Src-as1 (T338G) with the orthogonal ADP analog, N^6-(benzyl) ADP. This crystal structure of the remodeled protein-ligand interface is the first reported X-ray structure of a mutant kinase with a bound unnatural nucleotide analog. Comparison of the c-Src-as1 crystal structure with that of c-Src suggests that the effect of the active site mutation is confined to the adenine binding pocket and does not alter the protein (phospho-acceptor) binding site of the kinase.

To comprehensively compare the peptide specificity of wild-type and mutant kinases, we employed a degenerate peptide library of 50,625 putative peptide substrates as a specificity probe [14]. This degenerate peptide library technique provides a "fingerprint" of amino acid preferences for a given kinase, and it therefore defines a kinase's peptide specificity. One can compare these specificity fingerprints on a residue-by-residue basis to assess the phospho-acceptor substrate preferences of any kinase. Several analog-sensitive kinases were profiled and shown to have peptide specificity profiles identical to those of their wild-type counterparts. If any global conformational changes that drastically altered the substrate specificity of the engineered kinases had taken place, it would preclude the use of these mutants to identify the bona fide natural substrates of the wild-type enzymes. To further test the

[4] Correspondence: shokat@cmp.ucsf.edu
[5] Present address: Grand Valley State University, Department of Chemistry, Allendale, MI 49401.
[6] Present address: Genomics Institute of the Novartis Research Foundation, 3115 Merryfield Row, San Diego, CA 92121.
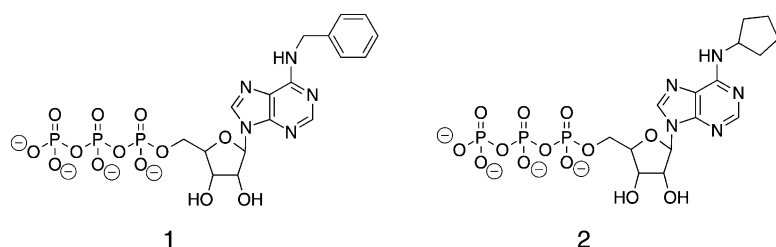
Figure 1. Chemical Structures of A*TP Analogs Used in This Study

1: N[6]-(benzyl) ATP; **2:** N[6]-(cyclopentyl) ATP. Definitions of analog-sensitive (as) kinase mutants used in this study.

Analog Sensitive Kinases:

v-Src-as1:	v-Src (I338G)
v-Src-as2:	v-Src (I338A)
Fyn-as1:	Fyn (T339G)

specificity of peptide phosphorylation by wild-type and analog-specific kinases, we kinetically assessed a panel of 11 optimal peptide substrates by determining the k_{cat} and K_M values for each peptide. No significant differences in kinetic parameters were observed between wild-type or analog-specific kinases. Since the analog-sensitive kinases have retained the peptide and protein substrate specificity of the wild-type enzymes, the mutant kinase/unnatural ATP system is a viable method for labeling the bona fide direct substrates of protein kinases [8].

Results and Discussion

Structural Analysis of c-Src-as1

In order to assess the effect of the mutation in the nucleotide binding pocket on the remainder of the kinase structure and to determine the binding mode of the ATP analog, we cocrystallized c-Src-as1 with the orthogonal substrate N[6]-(benzyl) ADP. The structure was determined at 2.8 Å resolution. The mutant kinase domain superimposes on wild-type c-Src (PDB ID 2SRC [40]) with an rms deviation of 0.35 Å for Cα atoms (Figure 2A). This difference is quite small given that the present structure was determined in a crystal lattice different than that of 2SRC. Thirty residues immediately surrounding the nucleotide binding pocket superimpose even more precisely (rmsd = 0.29 Å for Cα atoms). This superposition aligns the atoms in the adenine moiety of N[6]-(benzyl) ADP with the corresponding atoms in the AMP-PNP in 2SRC with a mean divergence of only 0.25 Å, which is within the expected experimental error of the present crystallographic analysis. The adenine and ribose portions of N[6]-(benzyl) ADP make the same hydrogen bonds observed in the wild-type structure; these bonds include those of the N[6] and N[1] nitrogens with the backbone carbonyl of Glu 339 and the amide of Met 341. Furthermore, we do not observe any side chain rearrangements in the nucleotide binding region of the mutant structure; the presence of the N[6]-(benzyl) group appears to be fully accommodated by the T338G mutation. The aromatic ring packs closely against the nearly planar engineered glycine residue (Figure 2B). The N[6]-(benzyl) ADP analog is orthogonal, i.e., not accepted by wild-type kinases, because of the severe steric clash created with the wild-type 338 residue (isoleucine

in v-Src, threonine in c-Src; Figure 2C). Interestingly, the benzyl substituent does not significantly intrude into a large existing pocket in c-Src, as we had previously suggested [15].

It is important to note that we have crystallized c-Src in an inactive conformation in which the peptide substrate binding site is not formed [40]. In the inactive conformation, the activation loop adopts an inhibitory helical conformation that buries Tyr 416 (the activating phosphorylation site in the activation loop). Phosphorylation of Tyr 416 induces a rearrangement that restores the activity of the kinase and creates the substrate binding site. Thus, based upon the present structure alone we cannot formally rule out the possibility of mutation- or analog-induced conformational changes in the phospho-acceptor site. To date, active Src has not been crystallized, but the structure of the Src-family kinase Lck has been determined in the active conformation [16]. Additionally, the active insulin receptor tyrosine kinase has been crystallized in complex with a peptide substrate [17]. Our comparison of the present structure with these active kinase structures suggests that it is extremely unlikely that the T338G mutation propagates a conformational change to the peptide substrate recognition region. The rearrangements induced upon activation are largely remote from the site of the mutation, and they do not affect the binding orientation or interactions of the adenine portion of the nucleotide. Thus, introduction of the mutation and orthogonal analog is likely to be structurally neutral in the enzyme's active conformation as well. Future studies of the effect of the T338G mutation on the fine regulation of the activity of c-Src will require kinetic examination of tail-phosphorylated and non-tail-phosphorylated forms of c-Src, as described by Miller and coworkers [18].

Combinatorial Peptide Library Design

The degenerate peptide library used here allows for the wild-type and mutant kinase specificity to be assessed with a chemically modified ATP, such as N[6]-(benzyl) ATP. This was not possible in the previous in vivo (cellular) test of kinase specificity, for which only ATP was present [9]. Peptide libraries have been used previously in order to probe the specificity of kinases, to determine kinase activity, and to predict the possible natural targets of kinases. Cantley and others have used peptide
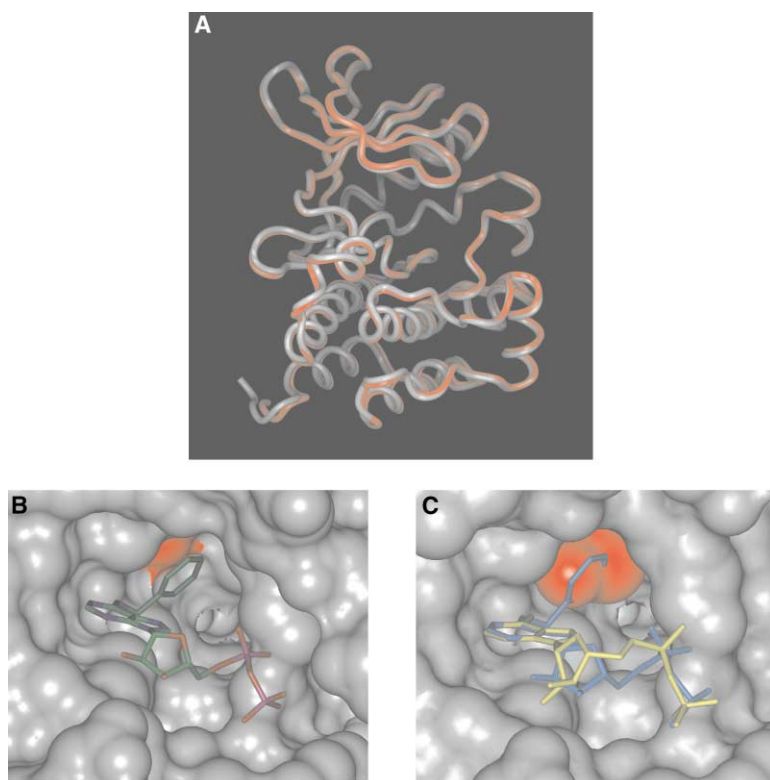
Figure 2. Comparison of Wild-Type and Analog-Specific c-Src Crystal Structures

(A) The crystal structure of c-Src-as1 superimposed on wild-type c-Src. c-Src-as1 is shown in gray, and c-Src is in red. The rmsd for the overlay is 0.35 Å.

(B) The binding of the A*TP analog, N^6-(benzyl) ADP to the mutant c-Src (T338G) kinase. The surface corresponding to the glycine residue at the 338 position is colored red. The benzyl ring of the A*TP analog projects into a pocket in the nucleotide binding cleft. This pocket is made accessible by the c-Src (T338G) point mutation. For clarity, the 11 residues that bind over the nucleotide at the front of the nucleotide cleft are omitted from the figure in order to more clearly show the surface at the back of the nucleotide binding pocket where the 338 residue lies. The omitted residues are c-Src 272–282.

(C) The steric clash of the wild-type c-Src threonine residue at the 338 position, shown in red, with the N^6-(benzyl) ATP analog (blue). The gray surface was built over the crystal structure of the mutant kinase overlayed with the wild-type c-Src crystal structure, and the surface was rendered over threonine 338 (red). The N^6-(benzyl) ADP (blue) is superimposed on the AMP-PNP ligand (yellow).

mixtures to determine optimal consensus sequences for members of both the serine/threonine and tyrosine kinase families [7, 14]. These libraries provide a probabilistic pattern for a given kinase's preference for amino acids at each position in the peptide chain, but they do not provide explicit substrate sequences. Alternative approaches, including the "Selectide" method of Lam et al.; the "SPOT" method, which involves immobilization of peptides in an array on cellulose paper; and the use of phage display peptide libraries, have also been developed [19–24]. These methods reveal explicit peptide sequences that serve as excellent kinase substrates. In many cases a "consensus sequence" can be inferred from the individual peptide sequence "hits."

For our studies, we chose the method developed by Cantley and coworkers because each preferred amino acid position in a peptide is measured in a semi-independent manner; this technique is therefore better for providing a specificity "fingerprint" of one kinase and then comparing that profile to that of another (modified) kinase. Methods that yield discrete sequences do not allow for adequate statistical comparisons between two different kinases. Finally, since many kinases have very loose specificity at certain amino acid positions, the degenerate peptide library approach is ideal for validation of "poor positional selectivity" in the same experiment in which highly preferred positions are also determined.

A peptide library was designed to probe two well-characterized features of c-Src target specificity. First, c-Src exhibits a high preference for the β-branched amino acid isoleucine (I) at the position immediately N-terminal (X) to the phosphorylated tyrosine [7]. Second, at the three adjacent N-terminal positions, (X), c-Src displays a *lack* of amino acid selectivity. With these two features in mind, we chose the peptide library sequence A-X-X-X-X-Y-G-E-F-K-K-K, where the positions marked X have been randomized to include the following 15 amino acids: A, E, D, F, G, H, I, K, L, M, N, P, Q, R, and V (tyrosine, threonine and serine were excluded to allow only one potential phosphorylation site, whereas tryptophan and cysteine were excluded because of known problems with oxidation during sequencing. The degenerate library therefore contains 15^4, or 50,625, possible unique peptide substrates. The polylysine tail was added to allow for binding of the peptides to phosphocellulose paper, which provides a means for washing away the negatively charged [γ-^{32}P] ATP and counting sample aliquots via liquid scintillation counting). One potential limitation of our library could be a bias against adjacent sterically bulky amino acids, but this limitation could be avoided by the use of a split-pool synthetic strategy [25]. The residues on the C-terminal side of tyrosine were not randomized but were kept constant as the sequence G-E-F-K-K-K (see below). The sequence G-E-F was determined to be a preferred substrate sequence for Src in previous work by Songyang et al. [7], so these residues were set, or "locked-in," to provide an enriched library of substrate peptides. This is desirable because some of the mutant kinases are less active than the wild-type enzymes and enough peptides must be phosphorylated in the library to provide sufficient sample amounts for sequence analysis [8, 9]. Our efforts to use a more highly degenerate peptide library, MAXXXXYXXXXAKKK, gave results that were difficult to interpret due to unphosphorylated tyrosine peptide contamination in the final purification step. (In attempts
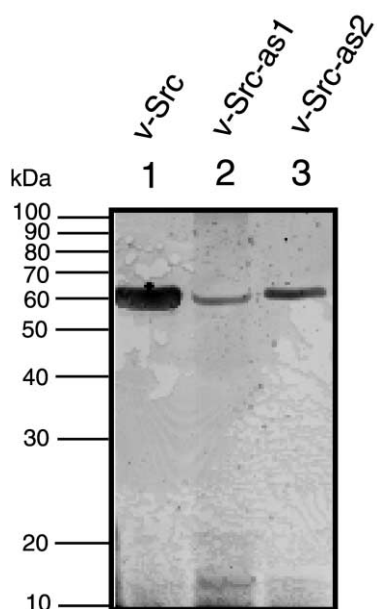
Figure 3. Silver-Stain SDS-PAGE (12%) of the v-Src Kinases Used in This Study

Lane 1, wild-type GST-v-Src (XD-4 construct: XD-4 [truncated v-Src kinase including the full catalytic domain, 77–225]); lane 2, GST-v-Src-as1; and lane 3, GST-v-Src-as2.

to rectify this problem, gallium metal IDA chelation columns were used based on the increased selectivity by gallium over iron for phosphate chelation [49]. Some enrichment, or improved retention, of phosphotyrosine was observed with the Ga-IDA column when compared to the Fe-IDA column, but this improvement was still not sufficient to allow assessment of the substrate specificity of mutant kinases.)

Peptide Library Phosphorylation
The peptide library A-X-X-X-X-Y-G-E-F-K-K-K was used to screen for the phospho-acceptor specificity profiles of v-Src and v-Src-as2 kinases with ATP and N^6-(benzyl) ATP, respectively. The v-Src-as2/ N^6-(benzyl) ATP system is the most *active and orthogonal* kinase-substrate pair we have identified to date [9]. The wild-type and mutant tyrosine kinases were purified as described in the Experimental Procedures section. An SDS-PAGE gel of the kinases used in these peptide library phosphorylation studies is shown in Figure 3. The peptide library was allowed to react with the kinase of interest to approximately 0.5%–1.0% conversion. If a kinase is allowed to react for extended periods of time with the peptide library, a leveling effect occurs as the most highly preferred amino acids are depleted at each position, thereby forcing the kinase to phosphorylate suboptimal and even poor substrates [14].

The purified peptide substrates were sequenced and analyzed for positional specificity (the average values of 2–3 separate trials were calculated). The amino acid distribution at each of the random positions was compared to that of the unreacted library to determine the enzymes' substrate specificity profile, or "fingerprint." This selectivity value was normalized to 15 because
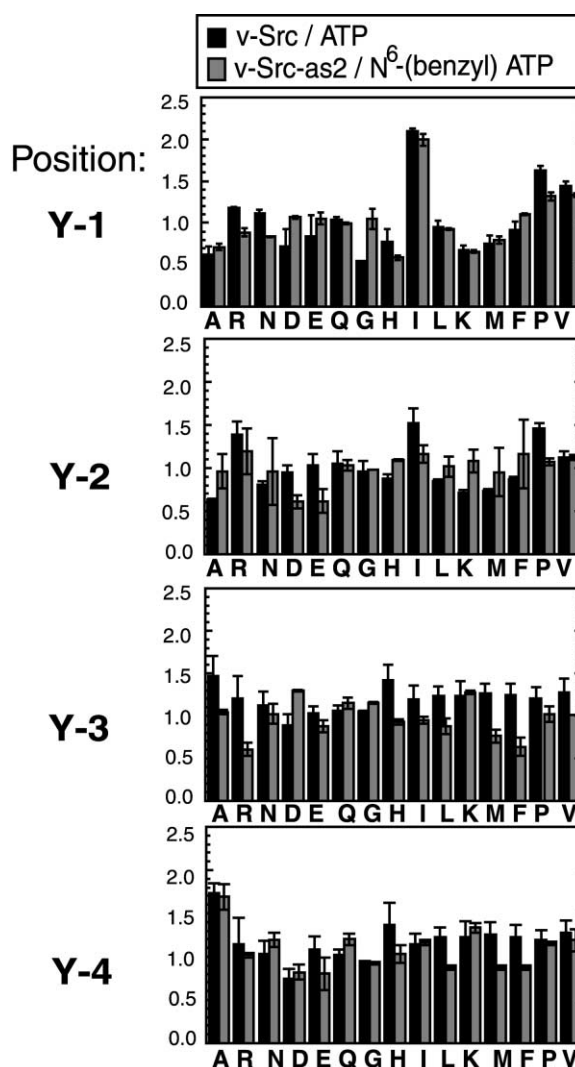

Figure 4. Specificity Profile of Wild-Type and Analog-Specific v-Src with N^6-(Benzyl) ATP

The specificity profile of the v-Src-as2 mutant with N^6-(benzyl) ATP is shown in gray. It is compared to the wild-type v-Src/ATP system, shown in black.

the enzyme has 15 different amino acids from which to choose at each position. A specificity value of 15 therefore would indicate that the kinase exclusively phosphorylated sequences with a particular amino acid at a given position. Any selectivity factor above 1.5 is significant for the tyrosine kinases [7, 14].

The selectivity profile comparison of wild-type v-Src/ATP and v-Src-as2/ N^6-(benzyl) ATP is shown in Figure 4. The Y-1 position, X, which corresponds to the residue closest to the phosphorylated tyrosine, showed an equivalent preference for isoleucine with both v-Src and the mutant v-Src-as2. The remaining randomized positions, Y-2, Y-3, and Y-4, showed no substantial changes in overall specificity between the wild-type and analog-specific kinases. Importantly, any specificity discrepancies between mutated and wild-type kinases occurred in the specificity range below 1.5, which represents a random choice of amino acids and is therefore not signif-

Table 1. Kinetic Values for Individual Peptide Substrates with Wild-Type and Mutant v-Src Kinases

Peptides	v-Src			v-Src-as2		
	K_M	k_{cat}	k_{cat}/K_M	K_M	k_{cat}	k_{cat}/K_M
1. RRLIEDAEYAARG	1200	160	0.1	1340	180	0.1
2. IYGEFKKK	65	1433	22.1	40	1137	28.4
3. EIYGEFKKK	45	1553	34.5	60	1795	29.9
4. IYGEWKKK	59	113	1.9	90	278	3.1
5. EIYWEFKKK	42	469	11.2	31	573	18.5
6. EIYWEWKKK	22	170	7.7	55	398	7.2
7. IYGKFKKK	120	196	1.6	141	245	1.7
8. NIYGHFKKK	346	140	0.4	228	129	0.6
9. AEIYAAPFAKKK	475	80	0.2	215	44	0.2
10. YIYGSFKKK	46	820	17.8	59	969	16.4
11. ESLYWSWPKKK	99	133	1.3	140	233	1.7

Comparison of peptide substrates of wild-type and mutant v-Src kinases. ATP is used as the phospho-donor. Wild-type v-Src is GST XD-4. The mutant v-Src-as2 is GST XD-4 (I338A). The enzyme concentration used in these assays was 1.5 nM for both the wild-type and the mutant kinases. K_M is in μM, k_{cat} is in min^{-1} and k_{cat}/K_M is min^{-1} μM^{-1}.

icant in this assay. For example, the difference between arginine specificity for wild-type v-Src and v-Src-as1 (1.3 versus 0.5) at the Y-3 position is not significant because both kinases show "no specificity" (both values are below the 1.5 cutoff) for arginine over any other amino acid residue. Cantley and others have adopted a similar cutoff value in kinase substrate library assays [7, 14]. Since the tyrosine kinases in general do not exhibit high sequence specificity, it will be important to confirm these findings with a ser/thr "as" kinase that exhibits high preference for a particular peptide sequence.

Kinase Peptide Specificity
As a complementary approach to combinatorial methods of assessing peptide specificity, we measured the kinetic parameters for phosphorylation of eleven individual peptides by v-Src and v-Src-as2. The peptides chosen contained different sequences C-terminal to the phosphorylatable position as well as amino acids that were not included in the random library for stability reasons. We chose peptide sequences (Table 1) known to be good substrates for tyrosine kinases [14, 26–29]. The RR-Src peptide sequence (RRLIEDAEYAARG) is derived from the consensus autophosphorylation site of c-Src [26, 30, 31]. The EIYGEFKKK and IYGEFKKK peptide substrates were determined from combinatorial peptide library studies by Songyang et al. [7]. The peptides IYGEWKKK and IYGKFKKK were synthesized to investigate the effect of size (Y + 3) and charge (Y + 2) modifications at several positions in the peptide. The IYGEWKKK, EIYWEFKKK, and EIYWEWKKK substrate peptides were chosen because the amino acid tryptophan was excluded from the peptide library studies due to problems with oxidation of the indole ring under Edman degradation reaction conditions. The peptide NIYGHFKKK was found to be a Src substrate when it was added to the tail of green fluorescent protein [32]. The Abl substrate AEIYAAPFAKKK was determined by the degenerate peptide library method to be the optimal

consensus sequence for Abl tyrosine kinase (a tyrosine kinase that is not in the Src family) and is used here as a nonoptimal substrate control for v-Src [7]. The ESLYWSWPKKK peptide was chosen based on the phage display experiments of Schmitz et al., who identified peptides of similar sequence (EESLYWSWPA) in two individual clones by c-Src phosphorylation after four cycles of selection [29]. The K_M and k_{cat} values of these peptides are shown in Table 1. v-Src and v-Src-as2 kinases exhibit similar K_M and k_{cat} values for all individual peptide substrates tested when ATP is the phospho-donor. These data strongly suggest that v-Src-as2 exhibits the same affinities and specificity constants as v-Src for a wide variety of peptide substrates.

Kinetic data were also collected for v-Src-as2 with N^6-(benzyl) ATP as the phospho-donor and for v-Src with ATP using the IYGEFKKK peptide substrate and were reported previously [9]. The kinetic constants for $[\gamma\text{-}^{32}P]$ ATP as the phospho-donor with v-Src are as follows: $k_{cat} = 132 \pm 4$ min^{-1}, $K_M = 11.5 \pm 1$ μM, $k_{cat}/K_M = 1.15 \times 10^7$ M^{-1} min^{-1}. The kinetic values with $[\gamma\text{-}^{32}P]$ N^6-(benzyl) ATP and v-Src-as2 are as follows: $k_{cat} = 65 \pm 2$ min^{-1}, $K_M = 20 \pm 3$ μM, $k_{cat}/K_M = 3.25 \times 10^6$ M^{-1} min^{-1} [9]. These values highlight the fact that the analog-specific allele of v-Src does not catalyze phosphorylation reactions with the orthogonal substrate, N^6-(benzyl) ATP, as well as it does with the natural substrate, ATP. Since the direct labeling of kinase substrates is carried out with $[\gamma\text{-}^{32}P]$ N^6-(benzyl) ATP as a tracer only, it is not essential that the k_{cat} values of ATP and $[\gamma\text{-}^{32}P]$ N^6-(benzyl) ATP match identically. More importantly, the analog-specific kinase does not exhibit any different peptide specificity than the unmutated wild-type kinase.

Peptide Specificity with Smaller Orthogonal Nucleotides
We also asked if a change in peptide substrate specificity might occur if the mutant kinase's enlarged binding pocket was less completely filled by the A*TP analog. To answer this question, we tested a less-flexible N^6 substituent with a shorter sidechain, $[\gamma\text{-}^{32}P]$ N^6-(cyclopentyl) ATP, **2**, with a kinase completely lacking a 338 side chain, v-Src-as1 (I338G) in the combinatorial peptide library assay. The specificity of v-Src-as1 was assessed with **2** since it models the largest possible newly created binding pocket we have described for an orthogonal nucleotide and mutant kinase [9]. The selectivity profile comparison of v-Src/ATP and v-Src-as1/ N^6-(cyclopentyl) ATP is shown in Figure 5. The Y-1 position, $\underline{X}$, that residue closest to the phosphorylated tyrosine, showed an equivalent preference for isoleucine for both the wild-type v-Src and the mutant v-Src-as1. The remaining randomized positions, Y-2, Y-3, and Y-4, showed no changes in specificity between the wild-type and mutant kinases. These data demonstrate that even a small A*TP analog used in conjunction with a mutation giving rise to access to a large complimentary binding pocket at the 338 position does not lead to any deleterious effects in the peptide specificity of the mutant enzyme/A*TP tagging system.

Peptide Specificity Profile for Fyn-as1
To determine if the A*TP/mutant chemical tagging system is extendable to other members of the Src family
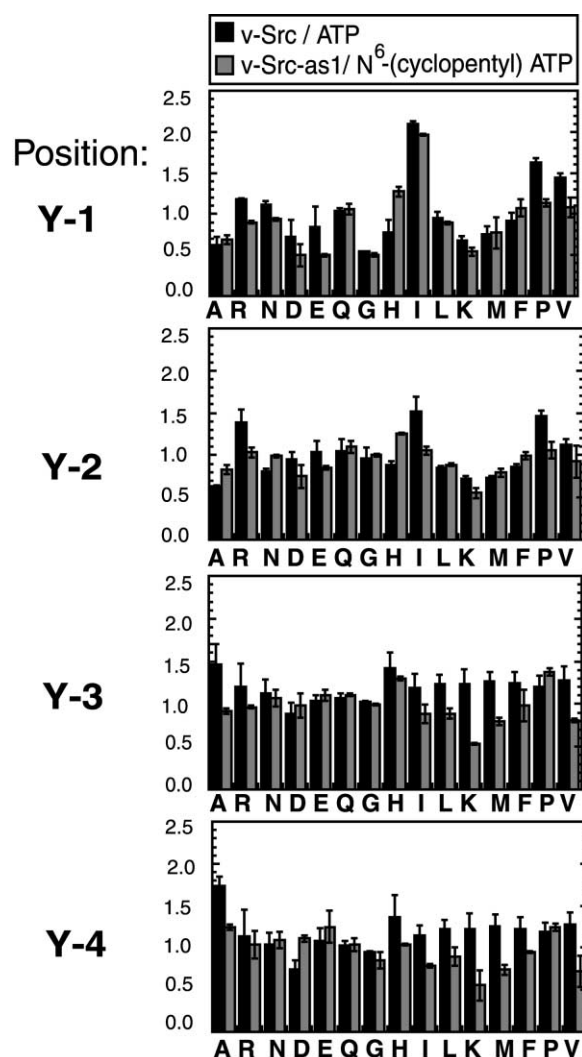
Figure 5. Specificity Profile of Wild-Type and Analog-Specific v-Src with N⁶-(Cyclopentyl) ATP

The specificity profile of the v-Src-as1 mutant with the A*TP analog N⁶-(cyclopentyl) ATP is shown in gray. It is compared to the wild-type v-Src/ATP system, shown in black.

Figure 6. The Specificity Profile of the Src Family Kinase Fyn.

The mutant Fyn-as1 tested with N⁶-(benzyl) ATP is shown in gray. It is compared to the wild-type Fyn/ATP system, shown in black.

of protein kinases without inducing a change in their peptide selectivity, we profiled the as1 allele of the tyrosine kinase Fyn. Wild-type full-length Fyn was reacted with ATP, and the specificity profile was compared to that of the space-creating mutant Fyn-as1 (T339G), with [γ-³²P] N⁶-(benzyl) ATP serving as the phospho-donor. The preference for isoleucine with the wild-type Fyn kinase at the Y-1 position (X) was similar to that found with wild-type v-Src (Figure 6). The mutant Fyn-as1/[γ-³²P] N⁶-(benzyl) ATP system displayed the same peptide specificity profile as wild-type Fyn.

Significance

Structural and functional analysis of the engineered kinase/modified ligand (A*TP) system confirms that no global conformational changes have taken place in the muta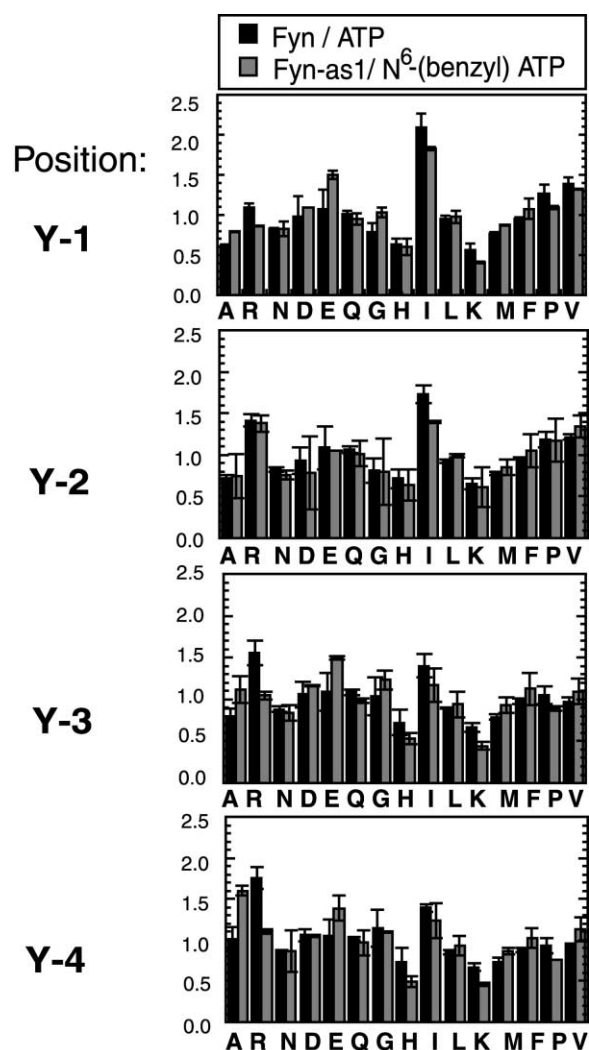nt kinase that would preclude its use as a substrate tagging tool. The engineered kinase, c-Src-as1, with N⁶-(benzyl) ADP bound and the wild-type c-Src with AMP-PNP bound show high structural similarity when they are superimposed. To assess the functional characteristics of the mutant kinase/A*TP system, we used a set of optimal Src peptide substrates and a peptide library of 50,625 possible tyrosine kinase substrates to probe the specificity of the wild-type and mutant kinases. The phospho-acceptor specificity of v-Src kinase, wild-type Fyn kinase, and several space-creating Src family mutants was determined. Several unnatural A*TP analogs were used as the phospho-donors for the mutant kinases, whereas ATP was used for the wild-type enzymes. The phospho-acceptor specificity profiles of the mutant enzymes with orthogonal A*TPs were found to be similar to the wild-type kinases, both in terms of amino acid residues that are highly selected for, and also in the preservation of low amino acid preferences in the

Table 2. Data Collection and Refinement Statistics

Resolution (Å)	2.80
Space group	$P2_12_12_1$
Unit cell (Å)	50.81, 87.68, 106.28
Molecules/a.s.u.	1
R_{sym} (%)	7.1
Reflections (total/unique)	46,549/12,715
Completeness (%)	98.6

Refinement Statistics	
Resolution Range (Å)	20.0–2.80
Protein atoms	3612
N_6 benzyl ADP atoms	34
Water molecules	46
R_{cyst}/R_{free} (%)	23.1/29.4
Rmsd bond length/angles	0.013 Å/1.771°

Unit cell constants and other structure determination statistics for the T338G c-Src mutant protein with N^6-(benzyl) ADP bound.

more-distal positions. The unaltered structural features of the allele-sensitive mutant kinase c-Src (T338G) and the unchanged functional peptide specificity of the mutant kinases validates the use of bulky A*TP analogs and space-creating mutants as tools for determining the direct bona fide cellular substrates of kinases.

Experimental Procedures

Materials
The following reagents were purchased from the indicated suppliers: [γ-^{32}P] ATP (6000 Ci/mmol), NEN; glutathione agarose beads, Sigma; Fmoc (9-Fluorenylmethoxycarbonyl)-protected amino acids and Wang resin, NovaBiochem; peptide synthesis reagents, Applied Biosystems; IDA (immobilized iminodiacetic acid) beads, B-PER II cell lysis reagent, and Gelcode silver stain kit, Pierce; DEAE Sephacel beads, Pharmacia Biotech; phosphocellulose (p81) paper discs, Whatman; Centricon-10 concentrators, Amicon.

Protein Expression and Purification
Expression and purification of GST-tagged kinases was carried out as described previously [33]. In brief, the protein expression and purification were as follows: single colonies were picked and grown overnight at 37°C in a shaker, diluted to 250 ml, and grown to cellular density of $A_{600\,nm} = 0.5$–0.7 (approximate time = 2 hr). Protein expression was then induced with IPTG (0.1 mM) for 5 hr, and the culture was then stored overnight at 4°C. Cells were lysed by either French press (the cell pellet was resuspended in the following buffer before pressing: 25 mM Tris [pH 8.0], 1 mM EDTA, 0.1 mM PMSF, 1 mM DTT). The GST fusion proteins were then purified on immobilized glutathione agarose beads, the beads were washed three times with 25 mM Tris (pH 8.0), 10 mM EDTA, and 100 mM NaCl, followed by two washes with 50 mM Tris (pH 8.0). The GST-fusion proteins were then eluted from the resin with free glutathione (freshly prepared: 5 mM reduced glutathione in 50 mM Tris [pH 8.0]) and concentrated in Centricon-10 concentrators. The enzyme purity and concentration were determined via silver-stained SDS-PAGE and via the Bradford assay [34].

Synthesis of Individual Peptide Substrates
Peptides of defined sequence were synthesized on an Applied Biosystems ABI431A automatic solid-phase peptide synthesizer by the use of Fmoc-protected amino acids, standard peptide synthesis protocols, and Wang polystyrene resin. *Fastmoc* reagent (HBTU [O-Benzotrizol-1-yl-N,N,N′-tetramethyluronium hexafluorophosphate] /HOBT [N-Hydroxybenzotriazole]) was used for coupling, and Reagent K was used for side chain deprotection and peptide cleavage from the solid support [35–37]. Isolation via ether precipitation afforded the individual peptides. Peptide sequences were determined via Edman degradation sequencing on an AB4768 protein sequencer [38]. Molecular weights were confirmed with electrospray mass spectrometry to observe the correct molecular ion for each of the peptides. Purity was assessed via reverse phase HPLC on a 300 Å wide pore 4.6 mm ID, 25 cm C18 analytical column with a gradient of 10%–100% acetonitrile in water (containing 0.1% TFA) at absorbances of 214 or 280 nm over 25 min (flow rate = 1.0 ml/min). Peptides tested were >98% pure as determined by HPLC.

Construction of Random Peptide Libraries
A degenerate library of tyrosine kinase peptide substrates, A-X-X-X-X-Y-G-E-F-K-K-K, was synthesized via solid-phase peptide synthesis on an ABI431A peptide synthesizer with Fmoc chemistry and HBTU/HOBT activation on Wang resin. The library was synthesized on a 1.2 mmol scale, and a 5 M excess of each amino acid (relative to resin substitution) was used for coupling. At the degenerate (X) positions, 15 amino acids (totaling a 5 M excess relative to resin substitution) were activated simultaneously with HOBT/HBTU and presented to the peptide resin. Sequencing of the final product confirmed that all 15 amino acids had indeed coupled within a factor of 3.5 of each other at each randomized position. The peptide library was cleaved from the solid support and were side chain deprotected with cleavage Reagent K. After ether precipitation, the library was dried to a fine white powder and stored at 4°C. The amino acid position denoted by the underline (X) indicates the degenerate position immediately N-terminal to the phosphorylatable tyrosine, whereas those positions not underlined (X) correspond to the more distant residues.

Synthesis of [γ-^{32}P] N^6-(Benzyl) ATP
N^6-(benzyl) adenosine was synthesized by refluxing 6-chloropurine riboside (1 mmol) (Aldrich) with benzylamine (5 mmol) in ethanol (10 ml) overnight. Ethanol was removed in vacuo, and the resulting oily residue obtained was crystallized from ethanol (yield 90%). N^6-(benzyl) ADP was synthesized according to the method of Hecht and Kozarich [39]. To an ice-cooled suspension of N^6-(benzyl) adenosine (68 mg, 0.2 mmol) in trimethyl phosphate (0.5 mmol), POCl$_3$ (0.025 ml) was added, and the reaction mixture was stirred at 0°C for 1 hr, after which the reaction was quenched with 5 ml of 1 M triethylammonium bicarbonate (TEAB buffer [pH 7.5]). Solvent was removed in vacuo at <40°C by rotary evaporation. The resulting slurry was purified on DEAE (A-25) Sephadex (Pharmacia) column with TEAB [pH 7.5] (0.1–0.5 M gradient). The purified N^6-(benzyl) AMP shows a retention time of 7.5 min on a strong anion exchange HPLC column (SAX, catalog # 83-E03-ETI, Varian) with a gradient of 5–750 mM ammonium phosphate (pH 3.9) in 10 min at a flow rate of 0.5 ml/min.

In the second step, a solution of N^6-(benzyl) AMP (44 mg, 0.1 mmol) and carbonyl diimidazole (81 mg, 0.5 mmol) in DMF (5 mL) was stirred at room temperature for 20 hr, after which methanol (35 μl) was added. After 1 hr, a solution of tributyl ammonium phosphate (1 mmol) was added in DMF (1 ml). The reaction was stirred for an additional 24 hr. After quenching the reaction mixture with 2 ml of TEAB buffer (pH 7.5), solvent was removed in vacuo at <40°C, and the residue was purified as described above (HPLC retention time 9.7 min).

N^6-(benzyl) ADP (2.5 μmol, molar absorbancy (ϵ_{max}) 15.4 × 10^3 at 265 nm at pH 7.00) was dissolved in DMF (200 μl), and carbonyl diimidazole (8 mg, 10 μmol) was added to it. The reaction mixture was stirred for 24 hr at room temperature, after which methanol (4 μl) was added and the reaction was stirred for an additional 1 hr. [γ-^{32}P] orthophosphoric acid (5 mCi, 8500 Ci/mmol) was dried in vacuo, dissolved in DMF (100 μl), and added to the reaction mixture. After the mixture was stirred for 24 hr, DMF was removed in vacuo, and the radiolabeled analog was purified by ion-exchange chromatography with DEAE (0.1–1 M TEAB buffer [pH 7.5]) at a flow rate of 0.5 ml/min (column size = 1.5 inches). The purified product was concentrated in vacuo at <40°C by rotary evaporation. The concentrated triphosphate was redissolved in 200 μl water, and the concentration was determined by scintillation counting (yield 20%). The [γ-^{32}P] N^6-(benzyl) ATP was characterized by coinjection of the radiolabeled material with an authentic sample of N^6-(benzyl) ATP on a strong anion exchange-HPLC column (retention time 11.2 min) [8].

Kinetic Assays

Kinase reactions with individual peptides were performed in kinase reaction buffer (50 mM Tris [pH 8.0], 10 mM $MgCl_2$, 100 mM NaCl) at a concentration of 100 μM [γ-^{32}P] ATP (1000 cpm/pmol). ATP analogs were used at the same concentration and specific activity. The reactions were quenched after 30 min by spotting 25 μl of the 30 μl total reaction mixture onto phosphocellulose (p81) paper. The disks were washed for 30 min with 10% acetic acid, followed by three 10 min washes with 0.5% phosphoric acid and a final 1 min wash with acetone. Liquid scintillation counting was used to measure the extent of peptide phosphorylation (cpm).

Peptide Library Assay

A typical experimental procedure is outlined below: 1 mg of the peptide library (average MW = 1800 g/mol) was added to the kinase of interest (10 nM). [γ-^{32}P] nucleotide (approximately 1000 cpm/ pmol) was added to the kinase in reaction buffer (50 mM Tris [pH 8.0], 10 mM $MgCl_2$, 100 mM NaCl). The reaction was then allowed to proceed to 0.5%–1.0% completion. The reaction progress was monitored by the removal of 5 μl aliquots, which were spotted on p81 phosphocellulose discs and counted for radioactivity as described above.

Purification of Phosphorylated Substrates

The separation of the phosphorylated peptide substrates from the nonphosphorylated reactants was achieved through a series of affinity columns as described by Cantley and coworkers [14]. The enzyme was separated from the peptide mixture with a Centricon-10 (10,000 Da molecular-weight cutoff) size exclusion filter. Excess nucleotide was removed via adsorption to 1.0 ml DEAE-Sephacel (acetate form) and eluted with 10 ml of 30% acetic acid [40]. Separation of the phosphorylated peptides from the unphosphorylated peptides was carried out on a 1.0 ml Fe^{+3} IDA chelation column that chelates phosphate groups selectively over other anionic species [7, 14, 41]. The acidic amino acid residues Glu and Asp can also act as weak ligands for this column. This nonselective adsorption results in an enrichment of acidic residues in the specificity analysis. To correct for this, we ran control reactions in which a solution of just the unphosphorylated peptide library was loaded onto the Fe^{+3} IDA column, and the bound peptide pool was sequenced as described [14]. Therefore, any enrichment of acidic residues due to retention on the chelation column is corrected by subtracting the values for these "mock" trials from the peptide library kinase reaction runs. The desired phosphorylated substrate peptides were then eluted from the column with 5–8 ml of 500 mM NH_4HCO_3 (pH 8.0). Fractions (0.5 ml) were collected and assayed for radioactivity. The fractions containing ^{32}P-labeled peptides were combined, lyophilized, and sequenced.

Library Sequencing

The phosphorylated substrate peptides were then sequenced via standard peptide Edman degradation liquid sequencing on an Applied Biosystems 4768 protein sequencer. Peak heights (pmol amounts) were used for amino acid quantification. Average standard deviations from 2–3 independent experiments for each kinase/nucleotide pair were used to determine error values. The peptide library, A-X-X-X-X-Y-G-E-F-K-K-K, was synthesized only once, and the same library was used in each of the specificity assays. The synthesized library of peptides was sequenced prior to reaction with the protein kinase in order to ensure that the library had a significant allocation of each amino acid at every random (degenerate) position.

Expression, Purification, and Crystallization of c-Src-as1

For structural studies, a DNA fragment encoding the SH3, SH2, kinase, and tail regions of human *c-Src* was subcloned into the baculovirus transfer vector pVL1392 (residues 83–533; by convention we use residue numbering corresponding to chicken c-Src). The Thr338Gly mutation was incorporated by the Quickchange mutagenesis procedure (Stratagene), and recombinant baculovirus was prepared by cotransfecting Sf-9 insect cells with the transfer vector and Baculogold DNA (Pharmingen). For large-scale production, 10 liters of Sf-9 cells cultured in TNM-FH plus 10% fetal bovine serum were infected with amplified virus and harvested 72 hr postin-

fection. The c-Src-as1 protein, homogeneously phosphorylated on Tyr 527, was purified as previously described for the wild-type protein [42]. The purified protein was concentrated to 10–15 mg/ml in storage buffer (20 mM HEPES [pH 7.6], 100 mM NaCl, and 10 mM DTT). For crystallization of c-Src-as1 with N^6-(benzyl) ADP, 50 μl of protein solution was combined with 50 μl of 1 mM N^6-(benzyl) ADP in storage buffer. All crystals were grown in hanging drops at 22°C by combining 2 μl of the above solution with 1 μl of a well solution (50 mM PIPES [pH 6.5], 12% PEG4000, 10 mM DTT). For low-temperature data collection, crystals were briefly "dunked" in crystallization solution and 20% glycerol and frozen at −165°C. Unit cell constants and other structure determination statistics are given in Table 2. Diffraction data were recorded with a Princeton 2k CCD detector on the A1 beamline at CHESS. All diffraction data were integrated and scaled with the programs DENZO and SCALEPACK [43].

Structure Determination

The structure of c-Src-as1 was determined by molecular replacement with the program AmoRe; the structure of c-Src complexed with AMP-PNP served as a search model [44, 45]. Calculated rotation and translation solutions were unambiguous. The properly positioned wild-type model with AMP-PNP deleted was subjected to overall and four-body rigid-body refinement (SH2/tail, SH3/linker, kinase N lobe, and kinase C lobes as rigid units) with XPLOR [46]. Electron density corresponding to N^6-(benzyl) ADP was obvious in initial maps, and N^6-(benzyl) ADP was built with O [47]. After a round of positional and B factor refinement, water molecules were added with the aid of the program of ARP [48]. Crystallographic R values and stereochemical parameters are presented in Table 2.

References

1. Hunter, T. (1987). One thousand and one protein kinases. Cell *50*, 823–829.
2. Hunter, T. (1995). Protein kinases and phosphatases: the Yin and Yang of protein phosphorylation and signaling. Cell *80*, 225–236.
3. Hunter, T., and Plowman, G.D. (1997). The protein kinases of budding yeast: six score and more. Trends Biochem. Sci. *22*, 18–22.
4. Cohen, G.B., Reb, R., and Baltimore, D. (1995). Modular binding domains in signal transduction proteins. Cell *80*, 237–248.
5. Pawson, T., and Scott, J.D. (1997). Science *278*, 2075–2080.
6. Hunter, T. (2000). Signaling—2000 and beyond. Cell *100*, 113–127.
7. Songyang, Z., Carraway, K.L., 3rd, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., et al. (1995). Catalytic specificity of protein tyrosine kinases is critical for selective signaling. Nature *373*, 536–539.
8. Shah, K., Liu, Y., Deirmengian, C., and Shokat, K. (1997). Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. Proc. Natl. Acad. Sci. USA *94*, 3565–3570.
9. Liu, Y., Shah, K., Yang, F., Witucki, L., and Shokat, K.M. (1998). Engineering Src family protein kinases with unnatural nucleotide specificity. Chem. Biol. *5*, 91–101.
10. Ulrich, S.M., Buzko, O., Shah, K., and Shokat, K.M. (2000). To-

wards the engineering of an orthogonal protein kinase/nucleotide triphosphate pair. Tetrahedron *56*, 9495–9502.

11. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., et al. (2000). Unnatural ligands for engineered proteins: new tools for chemical genetics. Annu. Rev. Biophys. Biomol. Struct. *29*, 577–606.

12. Habelhah, H., Shah, K., Huang, L., Ostareck-Lederer, A., Burlingame, A.L., Shokat, K.M., Hentze, M.W., and Ronai, Z. (2001). ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. Nat. Cell Biol. *3*, 325–330.

13. Habelhah, H., Shah, K., Huang, L., Burlingame, A.L., Shokat, K.M., and Ronai, Z. (2001). Identification of new JNK substrate using ATP pocket mutant JNK and a corresponding ATP analogue. J. Biol. Chem. *276*, 18090–18095.

14. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnica-Worms, H., and Cantley, L.C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. Curr. Biol. *4*, 973–982.

15. Liu, Y., Shah, K., Yang, F., Witucki, L., and Shokat, K.M. (1998). A molecular gate which controls unnatural ATP analog recognition by the tyrosine kinase v-Src. Bioorg. Med. Chem. *6*, 1219–1226.

16. Yamaguchi, H., and Hendrickson, W.A. (1996). Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. Nature *384*, 484–489.

17. Hubbard, S.R. (1997). Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J. *16*, 5572–5581.

18. Porter, M., Schindler, T., Kuriyan, J., and Miller, W.T. (2000). Reciprocal regulation of Hck activity by phosphorylation of Tyr(527) and Tyr(416). Effect of introducing a high affinity intramolecular SH2 ligand. J. Biol. Chem. *275*, 2721–2726.

19. Tegge, W., and Dostmann, W.R.G. (1995). Determination of cyclic nucleotide-dependent protein kinase substrate specificity by the use of peptide libraries on cellulose paper. Biochemistry *34*, 10569–10577.

20. Toomik, R., Edlund, M., Ek, P., Obrink, B., and Engstrom, L. (1996). Simultaneously synthesized peptides on continuous cellulose membranes as substrates for protein kinases. Pept. Res. *9*, 6–11.

21. Tegge, W.J., and Frank, R. (1998). Analysis of protein kinase substrate specificity by the use of peptide libraries on cellulose paper (SPOT-method). Methods Mol. Biol. *87*, 99–106.

22. Sparks, A.B., Rider, J.E., and Kay, B.K. (1998). Mapping the specificity of SH3 domains with phage-displayed random-peptide libraries. Methods Mol. Biol. *84*, 87–103.

23. Rickles, R.J., Botfield, M.C., Weng, Z., Taylor, J.A., Green, O.M., Brugge, J.S., and Zoller, M.J. (1994). Identification of Src, Fyn, Lyn, PI3K, and Abl SH3 domain ligands using phage display libraries. EMBO J. *13*, 5598–5604.

24. Rickles, R.J., Botfield, M.C., Zhou, X.-M., Henry, P.A., Brugge, J.S., and Zoller, M.J. (1995). Phage display selection of ligand residues important for Src homology 3 domain binding specificity. Proc. Natl. Acad. Sci. USA *92*, 10,909–10,913.

25. Gast, R., Glokler, J., Hoxter, M., Kiess, M., Frank, R., and Tegge, W. (1999). Method for determining protein kinase substrate specificities by the phosphorylation of peptide libraries on beads, phosphate-specific staining, automated sorting, and sequencing. Anal. Biochem. *276*, 227–241.

26. Casnellie, J.E., Harrison, M.L., Pike, L.J., Hellstrom, K.E., and Krebs, E.G. (1982). Phosphorylation of synthetic peptides by a tyrosine protein kinase from the particulate fraction of a lymphoma cell line. Proc. Natl. Acad. Sci. USA *79*, 282–286.

27. Dente, L., Vertriani, C., Zucconi, A., Pelicci, G., Lanfrancone, L., Pelicci, P.G., and Cesareni, G. (1997). Modified phage peptide libraries as a tool to study specificity of phosphorylation and recognition of tyrosine containing peptides. J. Mol. Biol. *269*, 694–703.

28. Lam, K., Wu, J., and Lou, Q. (1995). Identification and characterization of a novel synthetic peptide substrate specific for Src-family protein tyrosine kinases. Int. J. Pept. Protein Res. *45*, 587–592.

29. Schmitz, R., Baumann, G., and Gram, H. (1996). Catalytic speci-ficity of phosphotyrosine kinases Blk, Lyn, c-Src and Syk as assessed by phage display. J. Mol. Biol. *260*, 664–677.

30. Czernilofsky, A.P., Levinson, A.D., Varmus, H.E., Bishop, J.M., Tischer, E., and Goodman, H.M. (1980). Nucleotide sequence of an avian sarcoma virus oncogene (Src) and proposed amino acid sequence for gene product. Nature *287*, 198–203.

31. Feder, D., and Bishop, J.M. (1990). J. Biol. Chem. *265*, 8205–8211.

32. Yang, F., Liu, Y., Bixby, S., Friedman, J., and Shokat, K. (1999). Highly efficient green fluorescent protein-based kinase substrates. Anal. Biochem. *266*, 167–173.

33. Liu, Y., Witucki, L.A., Shah, K., Bishop, A.C., and Shokat, K.M. (2000). Src-Abl tyrosine kinase chimeras: replacement of the adenine binding pocket of c-Abl with v-Src to swap nucleotide and inhibitor specificities. Biochemistry *39*, 14400–14408.

34. Bradford, M.M. (1976). Anal. Biochem. *72*, 248–254.

35. Fields, G.B., and Noble, R.L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. Int. J. Pept. Protein Res. *35*, 161.

36. King, D.S., Fields, C.G., and Fields, G.B. (1990). A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. Int. J. Pept. Protein Res. *36*, 255.

37. Wang, S.S. (1973). p-Alkoxybenzyl alcohol resin and p-alkoxybenzyloxycarbonyl hydrazide resin for solid phase synthesis of protected peptide fragments. J. Am. Chem. Soc. *95*, 1328–1333.

38. Edman, P. (1950). Method for the determination of the amino acid sequence in peptides. Acta Chem. Scand. *4*, 283.

39. Hecht, S.M., and Kozarich, J.W. (1973). A chemical synthesis of adenosine 5′ [γ-32P] triphosphate. Biochim. Biophys. Acta *331*, 307–309.

40. Kemp, B. (1976). Synthetic Hexapeptide Substrates and Inhibitors of 3′:5′-Cyclic AMP-Dependent Protein Kinase. Proc. Natl. Acad. Sci. USA *73*, 1038–1042.

41. Muszynska, G., Andersson, L., Porath, J. (1986). Selective adsorption of phosphoproteins on gel-immobilized ferric chelate. Biochemistry *25*, 6850–6853.

42. Xu, W., Harrison, S.C., and Eck, M.J. (1997). Three-dimensional structure of the tyrosine kinase c-Src. Nature *385*, 595–602.

43. Otwinowski, Z. (1993). Oscillation Data Reduction Program. (Daresbury, UK: SERC Daresbury Laboratory).

44. Navaza, J. (1992). AMoRe: A New Package for Molecular Replacement. (Daresbury, UK: SERC Daresbury Laboratory).

45. Xu, W., Doshi, A., Lei, M., Eck, M.J., and Harrison, S.C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol. Cell *3*, 629–638.

46. Brunger, A.T. (1992). X-PLOR, Version 3.0: A System for Crystallography and NMR. (New Haven, Connecticut: Yale University).

47. Jones, T.A., Bergdoll, M., and Kjeldgaard, M. (1989) Crystallographic Computing and Modeling Methods in Molecular Design (New York: Springer).

48. Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. Acta Crystallogr. D *49*, 129–147.

49. Posewitz, M.C., and Tempest, P. (1999). Immobilized gallium (III) affinity chromatography of phosphopeptides. Anal. Chem. *71*, 2883–2892.

Accession Numbers

Coordinates have been deposited with the Protein Data Bank, accession code 1KSW.