Articles

Mapping the Subsite Preferences of Protein Tyrosine Phosphatase PTP-1B Using Combinatorial Chemistry Approaches

Matthew C. Pellegrini, Hongbin Liang, Sreekala Mandiyan, Karen Wang, Anton Yuryev, Isidoros Vlattas, Teri Sytwu, Yu-Chin Li, and Lawrence P. Wennogle*

Novartis Pharmaceutical Corporation, 556 Morris Avenue, Summit, New Jersey 07901

Received June 16, 1998; Revised Manuscript Received August 26, 1998

ABSTRACT: Protein tyrosine phosphatases (PTPases) are important regulators of signal transduction systems, but the specificity of their action is largely unexplored. We have approached this problem by attempting to map the subsite preferences of these enzymes using combinatorial chemistry approaches. Protein-tyrosine peptidomimetics containing nonhydrolyzable phosphotyrosine analogues bind to PTPases with high affinity and act as competitive inhibitors of phosphatase activity. Human PTP-1B, a PTPase implicated to play an important role in the regulation of growth factor signal transduction pathways, was used to screen a synthetic combinatorial library containing malonyltyrosine as a phosphotyrosine mimic. Using two cross-validating combinatorial chemistry screening approaches, one using an iterative method and the other employing library affinity selection−mass spectrometric detection, peptides with high affinity for PTP-1B were identified and subsite preferences were detailed by quantitatively comparing residues of different character. Consistent with previous observations, acidic residues were preferred in subsites X_{-3} and X_{-2}. In contrast, aromatic substitutions were clearly preferred at the X_{-1} subsite. This information supports the concept that this class of enzymes may have high substrate specificity as dictated by the sequence proximal to the phosphorylation site. The results are discussed with regards to the use of combinatorial techniques in order to elucidate the interplay between enzyme subsites.

Protein-tyrosine phosphatases (PTPases)[1] comprise a large superfamily of enzymes that catalyzes the hydrolysis of phosphate from phosphotyrosine (pY) residues (*1*, *2*). This family of enzymes is characterized by having a structurally conserved catalytic domain of approximately 250 amino acid and a unique cysteine-containing signature sequence, I/VHCXAGXXRS/TG [additional domains may include src homology-2 (SH2), transmembrane, and extracellular domains]. PTPases function in maintaining the steady-state balance of protein-tyrosine phosphorylation in mammalian signal transduction pathways; in so doing, they regulate important biological functions including metabolism, proliferation, and differentiation. In addition, certain PTPases have been identified as being potent disease causing virulence factors (*3*−*5*).

Human PTP-1B is the best characterized member of the mammalian PTPase superfamily (*6*−*8*) and has been implicated in playing an important role in the regulation of growth factor signal transduction pathways such as the insulin receptor system (*9*−*14*). Following insulin binding, the insulin receptor undergoes autophosphorylation of tyrosine residues in the kinase activation loop and parallel activation of intrinsic tyrosine kinase activity toward endogenous substrates (*15*). Tyrosine autophosphorylation is essential for activity of the insulin receptor, and it is increasingly evident that PTP-1B and/or closely associated PTPases deactivate the insulin receptor by dephosphorylation [reviewed by Goldstein (*9*)].

A detailed understanding of the substrate specificity of mammalian PTPases is generally lacking. Kinetic studies, however, have provided some evidence for the ability of related PTPases to discriminate between substrates at the primary structural level (*16*−*21*). Important structural determinants of substrate recognition are included in short pY-containing peptides (*16*, *19*). Alanine scanning of the $EGFR_{988-998}$ phosphopeptide has suggested that substrate recognition is primarily dictated by subsites proximal and N-terminal to the phosphorylation site.

* Corresponding author. Phone: (908) 277-5697. Fax: (908) 277-4739. E-mail: lawrence.wennogle@pharma.novartis.com.

[1] Abbreviations: PBS, phosphate-buffered saline; MS, mass spectral analysis; ES-MS, electrospray MS; ES-MS/MS, electrospray-tandem MS; LAS-MS, library affinity selection−mass spectrometric detection; GST, glutathione-*S*-transferase; FMOC, *N*-(9-fluorenylmethoxycarbonyl) blocking group; amino acid abbreviations are Ala (A), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Pro (P), Gln(Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W), Tyr (Y), hydroxyPro (hP); pY, phosphotyrosine; mY, malonyl tyrosine; f2pmp-F, *p*-difluorophosphonate-phenylalanine; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; PTPase, protein tyrosine phosphatase; BSA, bovine serum albumin; TFA, trifluoroacetic acid; HPLC, high-pressure liquid chromatography; DMSO, dimethyl sulfoxide.

10.1021/bi981427+ CCC: $15.00
Published on Web 10/22/1998

We sought a more detailed understanding of the subsite preference of human PTP-1B to better understand substrate specificity and to potentially lead to a rational design of competitive inhibitors. Accordingly, a "synthetic combinatorial library" of protein-tyrosine peptidomimetics was designed based on the EGFR autophosphorylation site at Y_{992} ($EGFR_{988-995}$). Three residues proximal and N-terminal to the phosphorylation site were randomized using a "split-synthesis" approach (*22*). Split-synthesis permitted the construction of mixed sequence peptides containing the nonnatural amino acid malonyltyrosine (mY), a nonhydrolyzable pY mimic used to decouple binding affinity from catalytic activity (*23*, *24*). Two cross-validating screening approaches, an "iterative" approach (*25*−*27*), and a library affinity selection−mass spectrometric detection (LAS-MS) approach (*28*) were employed. Through this effort, protein-tyrosine peptidomimetics that function as potent inhibitors of PTP-1B phosphatase activity were identified, and a novel subsite preference for PTP-1B was documented.

METHODS AND MATERIALS

Peptide Synthesis. The phosphopeptide substrates were synthesized using Rink resin (Midwest Bio-Tech, Fishers, IN) and standard FMOC solid-phase peptide chemistry. The synthetic combinatorial library was prepared by split synthesis using a divide, couple, and recombine procedure with Rink resin and standard FMOC solid-phase peptide chemistry. Following deprotection, cleavage, and extraction, the peptides were purified using reversed-phase HPLC, lyophilized, resuspended in DMSO, and stored at 4 °C as 20 mM stock solutions. The library was defined as DXXX-(mY)LIP, where mY denotes malonyltyrosine, and X denotes an equal mixture of 19 standard L-amino-acids (cysteine excluded) plus L-hydroxyproline. Particular care was taken to ensure the integrity of various libraries. The split synthesis method allowed individual reactions to be individually monitored and therefore taken to completion and assured near equal contribution of individual building blocks at a each step in a library synthesis. As further confirmation, libraries were routinely evaluated by mass spectral (MS) analysis, as detailed below. For individual peptides synthesis, HPLC and amino acid analysis was performed to allow for accurate quantitation.

PTP-1B Purification. Glutathione-Sepharose beads were purchased from Pharmacia (Piscataway, NJ). All other chemicals were purchased from Sigma (St. Louis, MO). Human PTP-1B (amino acids 1−411, lacking the endogenous C-terminal 23 amino acids) was cloned by PCR from a human hippocampus cDNA library and inserted into pGEX5x-2 (Pharmacia) at the *Eco*R1 and *Sma*1 restriction endonuclease sites. *Escherichia coli* strain BL21 was transformed with the above clone and stored as a stock in 15% glycerol at −70 °C. The above construct allowed high-level, IPTG-inducible expression of PTP-1B fused at the N-terminus to glutathione-*S*-transferase. GST-PTP-1B preparations were grown at a 1 L scale to an OD_{590} of 0.6 at 37 °C, induced with 1.5 mM IPTG, and harvested 3 h later (all subsequent steps were performed at 4 °C). Cells were pelleted, washed with 100 mL of phosphate-buffered saline (PBS), and resuspended in 10 mL of PBS supplemented with complete protease inhibitor cocktail (Boehringer Mannheim) and 0.1% Triton X-100. The suspension was sonicated 5 × 30 s on ice using a Branson sonicator with a microtip and then centrifuged at 20 000 rpm for 20 min in a SS-34 rotor. The supernatant was decanted, incubated with 1 mL of glutathione-Sepharose beads (as a 50% slury of beads prewashed with PBS) for 1 h with gentle shaking, and then loaded onto a column. After 3 × 10 mL washes of the column with PBS, GST-PTP-1B fusion protein was eluted by treatment with 1 mL of 10 mM reduced glutathione in PBS. The preparation contained 2−5 mg of GST-PTP-1B/mL of beads as verified by protein analysis and by SDS−PAGE, after removal from beads. Enzyme was passed through a NAP column (Pharmacia) preequilibrated with 50 mM HEPES, pH 7.0, 1 mM DTT, concentrated with an Amicon YM-10 membrane, and stored at −70 °C until use.

Malachite Green PTPase Assay. PTP-1B activity was measured using a Malachite Green color reagent to detect release of inorganic phosphate from phosphopeptides. The Malachite Green reagent was prepared as a 3:1 mixture of 0.045% malachite green and 4.2% ammonium molybdate supplemented with 0.01% TritonX-100. Reaction mixtures (100 μL) containing GST-PTP-1B and phosphopeptide substrates in buffer A (50 mM Tris, pH 7.5, 50 mM NaCl, 3 mM DTT, and 0.1% BSA) were incubated for 15 min at 37 °C with gentle shaking and then quenched with 200 μL of Malachite Green reagent. After 30 min at room temperature with gentle shaking, the absorbance at 620 nanometers was measured using a Ceres UV900C microtiter plate spectrometer (Bio-Tek Instruments). Specific activity was defined as the difference between samples with and without substrate. Inhibition was measured as the reduction in the activity of PTP-1B (20 nM) with the phosphopeptide substrate GNGDPpYPMSPKS (30 μM) in the presence of inhibitor.

PTP-1B Library Affinity Selection. GST-PTP-1B fusion protein from a 1 l fermentation was immobilized onto 1 mL of glutathione-Sepharose beads as above and washed with 3 × 10 mL in buffer B (150 mM NaCl and 50 mM ammonium acetate, pH 6.6). For the batch method of library affinity selection, a 10 mL suspension of beads and library containing 50 nmol of immobilized GST-PTP-1B and 2.0 μmol of DDXX(mY)LIP in buffer B was used. (Since the library contained 400 components, 5.0 nmol of each component peptide was present.) Mixtures were incubated at 4 °C for 2 h with gentle shaking, washed with 3 × 5 mL of buffer B to remove unbound library components, and then washed with 3 × 5 mL of buffer C (500 mM ammonium acetate, pH 6.6) to remove weakly bound library components. Library components with high affinity for PTP-1B were eluted by washing with 3 × 5 mL of buffer D (100 mM triethylamine acetate, pH 3.5), and the eluted peptides were monitored by ES-MS and ES-MS/MS (see below).

Mass Spectrometry. ES-MS and ES-MS/MS were performed with a PE-SCIEX API III triple-quadruple mass spectrometer (Thornhill, Ontario) in the positive ion mode. All analyses were carried out via LC/MS. Peptides contained in the buffer D wash were lyophilized, reconstituted in 2% acetonitrile/0.1% TFA in water, and injected into a 1.0 × 150 mm Vydac C18 microbore column for on-line analysis. A gradient program of 2−2−50−50−90−90% B at 0−5−7−17−20−30 min was applied where A was 0.05% TFA in water and B was acetonitrile. The flow rate was 50 μL/min. ES-MS conditions utilized a spray voltage of 5 kV,

an orifice voltage of 60 V, a scan range of 900−1300 *m/z* units in 0.1 *m/z* increments, and a speed of 2 s/scan. Purified air was used as a nebulization gas (30 psi) and ultrapure nitrogen was used as a curtain gas at a flow rate of 1 L/min. ES-MS/MS experiments were carried out using ultrapure argon at $\sim 300 \times 10^{-12}$ molecules/cm^2 collision gas thickness, a 30 eV collision energy, a scan range of 50 (parent ion + ca. 10) *m/z* units in 0.33 *m/z* increments

RESULTS

Iterative Approach. The screening of synthetic peptide combinatorial libraries using an iterative approach permitted rapid and efficient structural deconvolution leading to optimized peptide ligands. To accommodate this process, a Malachite Green enzyme assay (*29*, *30*) was employed to measure PTP-1B catalysis in a high-throughput, solution format. Free inorganic phosphate liberated from the phosphopeptide substrate GNGDPpYPMSPKS by recombinant PTP-1B was detected by measuring OD_{620}. Control experiments demonstrated linearity of the enzyme with respect to both time and enzyme concentration (not shown), and saturation analysis revealed a K_m of 68 μM and V_{max} of 4 μmol/min/mg (not shown). The kinetic properties of the GST-PTP-1B fusion enzyme were similar to literature values of PTP-1B isolated from natural and recombinant sources and indicate that the activity correctly reflects the native enzyme (discussed by Barford et al. (*31*), Zhang and Dixon (*32*), and Milarski et al. (*18*).

The initial library synthesis was defined as $DO_{-3}X_{-2}X_{-1}$(mY)LIP, where O denotes a defined residue subsite having each of the 19 standard L-amino-acids (cysteine excluded) plus L-hydroxyproline, and X denotes a mixed residue subsite having equal fractions of these 20 building blocks (the standard PTPase nomenclature names residues N-terminal to pY as successively negative subsites). Library O_{-3} contained 20 separate sublibraries, each with a different fixed O_{-3} residue and each containing 400 different peptides (20^2). The 20 sublibraries were screened and rank ordered for their ability to inhibit PTP-1B catalysis to define the amino acid preference at the −3 subsite (Figure 1A). Of these, DDXX(mY)LIP caused the greatest inhibition. Other O_{-3} residues that provided strong inhibition (>50%) were G, W, hP, and E.

The second library synthesis was defined as $DD_{-3}O_{-2}X_{-1}$(mY)LIP, where D_{-3} corresponds to the most effective inhibitory residue from the first iteration. Library O_{-2} contained 20 separate sublibraries, each with a different fixed O_{-2} residue and each containing 20 different peptides. The 20 sublibraries were screened and rank ordered for their ability to inhibit PTP-1B catalysis to define the amino acid preference at the −2 subsite (Figure 1B). Of these, DDEX(mY)LIP caused the greatest inhibition. Other O_{-2} residues that provided strong inhibition were D and S, while K was the weakest substitution.

The third library synthesis was defined as $DD_{-3}E_{-2}O_{-1}$(mY)LIP, where $D_{-3}E_{-2}$ corresponds to the most effective inhibitory sequence from the second iteration. Library O_{-1} contained 20 separate peptides, each with a different fixed O_{-1} residue. Each of the 20 peptides was screened and rank ordered for their ability to inhibit PTP-1B catalysis to define the amino acid preference at the −1 subsite (Figure 1C). Of these, DDEW(mY)LIP caused the greatest inhibition of PTP-1B phosphatase activity. Importantly, aromatic residues were preferred (W, Y, and F) and strong discriminatory behavior was observed.

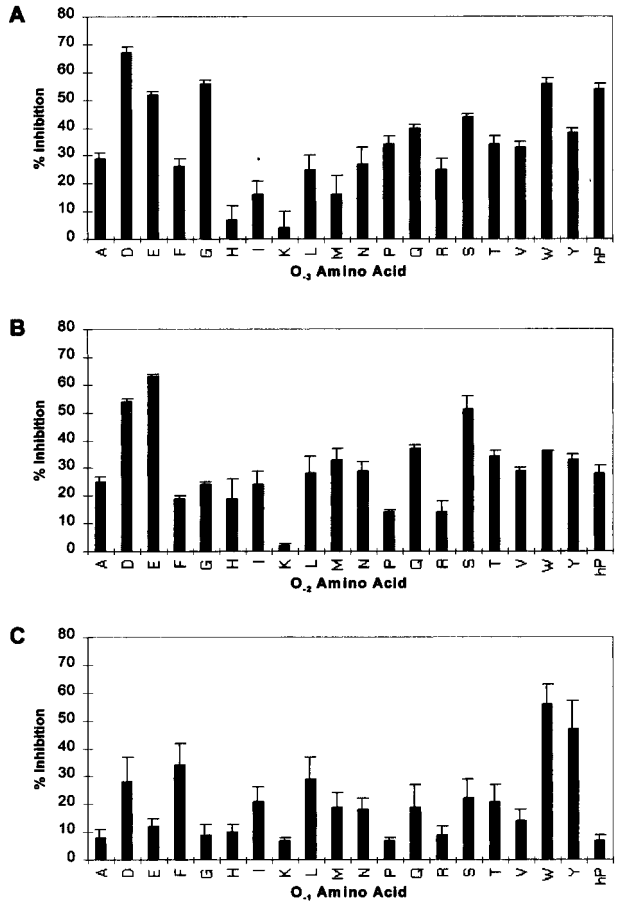

FIGURE 1: Iterative screening of a combinatorial library. The DXXX(mY)LIP library was screened for the ability to inhibit PTP-1B activity using then Malachite Green assay described in the Materials and Methods using (A) library O_{-3}; $DO_{-3}X_{-2}X_{-1}$(mY)LIP screened at a peptide concentration of 20 μM, (B) library O_{-2}; $DD_{-3}O_{-2}X_{-1}$(mY)LIP screened at 2 μM, and (C) library O_{-1}; $DD_{-3}E_{-2}O_{-1}$(mY)LIP screened at 0.2 μM. The results are the average (±SEM) of three independent experiments performed in triplicate.

Each step in the library iteration resulted in inhibitors with greater affinity for the PTP-1B. This can be seen from the inhibition data from Figure 1, where decreasing screening concentrations were used: 20, 2, and 0.2 μM, for respective rounds. When IC_{50} values were determined for the best inhibitors of each round, DDXX(mY)LIP had an IC_{50} of 8.6 (± 2.3) μM, DDEX(mY)LIP had an IC_{50} of 6.8 (± 0.7) μM, DDEW(mY)LIP had an IC_{50} of 0.73 (± 0.17) μM, and DDEY(mY)LIP had and IC50 of 0.76 (±0.22) μM. In general, basic residues were not favored and clear discriminatory behavior was demonstrated at each of the randomized subsites. Assuming normal competitive inhibition kinetics and a Hill coefficient of unity, the range of activities between individual pools within each iteration correlated to roughly a 100-fold difference in binding affinities.

The data from Figure 1 has been evaluated for significant differences using a one-way ANOVA (SigmaStat 2.0) with a Tukey Test for multiple comparisons. All pairwise comparisons were made to determine significance ($p < 0.05$)

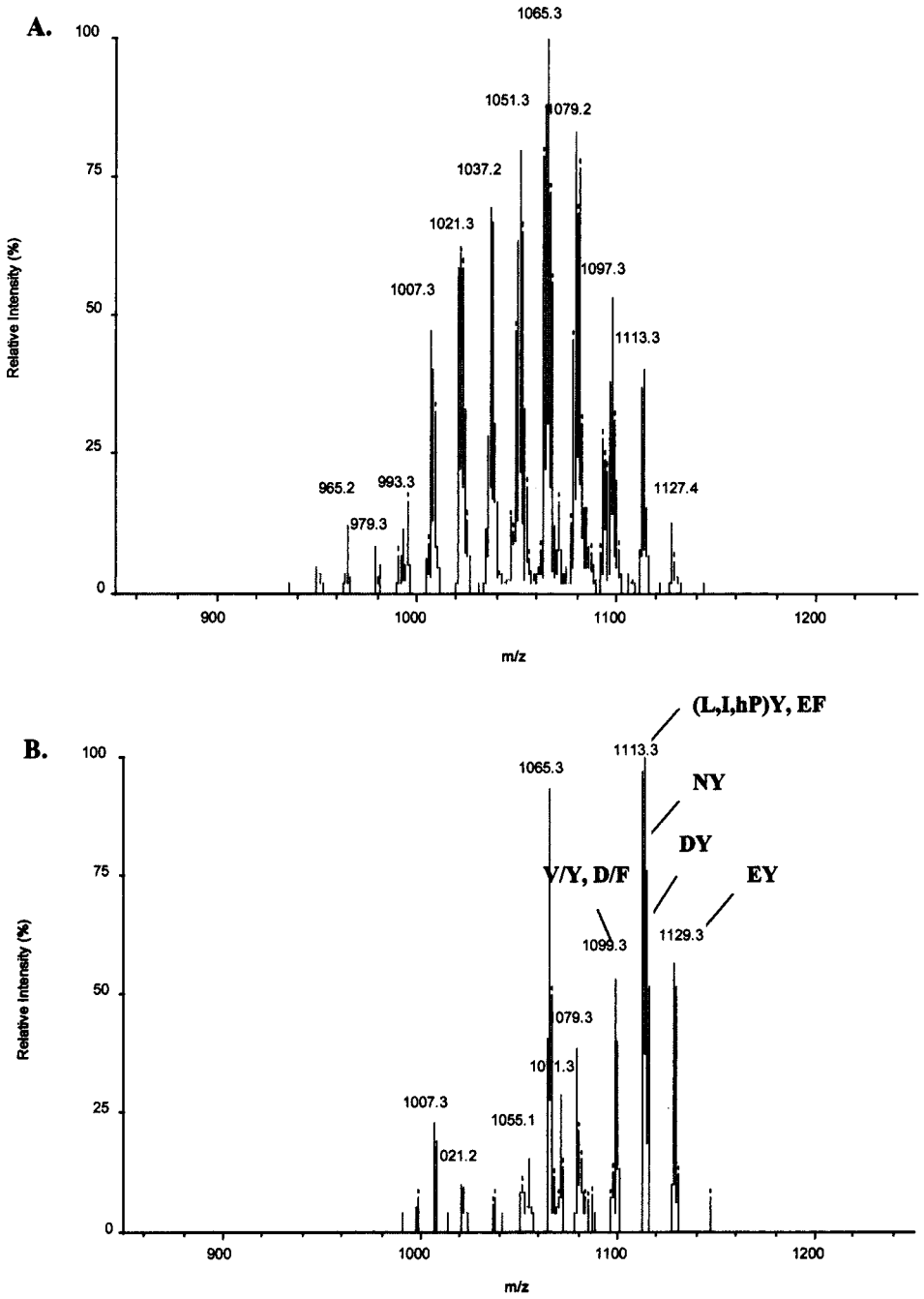

FIGURE 2: ES-MS spectra of DDXX(mY)LIP from library O_{-3} before and after library affinity selection. (A) ES-MS spectrum of the intact library. (B) ES-MS spectrum of the library affinity selected components. The protocol is detailed in the Materials and Methods. Weakly bound library components were removed with high salt (buffer C) and affinity selected components were eluted with low pH (buffer D). The eluted peptides were monitored by ES-MS with a scan range of 900−1300 *m*/*z* units. Amino acid designations are for the sequence $X_{-2}X_{-1}$ and are detailed in Table 1.

between different amino acids at each site. These evaluations extended the graphical representation presented in Figure 1. In site $X_{-1,}$ "W" was found significantly greater than all other groups except "F and Y". "Y" was significantly greater than all but D, F, L, and W. While "F" was ranked as the third largest inhibition, it was statistically different from "G, H, K, Q, S, and hP". These preferences between amino acids at a particular site resulted in some unexpected observations. One example is the lack of significance between "F, D and L" at X_{-1}, since these are very different amino acids. Another example was the finding that "G" was not significantly different from "E, D, and W" in X_{-3}. While it is difficult to reconcile these comparisons on theoretical grounds, it may attest to the remarkable selectivity of protein interactions and could support the dynamic fit into a malleable contact surface. Taking these interpretations one step further in the reverse sense, unfavorable residues are found at each subsite. Positive residues, particularly "K", showed low affinity at all three sites. Negative preferences, like the positive preferences, are likely to play a significant role in driving specific and high-affinity interactions.

LAS-MS Experiments. To strengthen the observations made with the iterative screening approach, selection of high-affinity sequences from the most effective inhibitory O_{-3} sublibrary was performed using library affinity selection with mass spectrometric detection (LAS-MS). Coupling an

affinity selection process with electrospray ionization mass spectrometry (ES-MS) and electrospray tandem mass spectrometry (ES-MS/MS) affords a powerful means for the structural deconvolution of optimized ligands from synthetic combinatorial libraries. Peak redundancies may exist because certain building blocks have identical mass. Furthermore, ES-MS will not discern between ligands with the same amino acid content but different sequence, for example the sequence DDEY(mY)LIP and DDYE(mY)LIP. Because ES-MS analysis alone will not resolve these redundancies, ES-MS/MS was employed due to its sequencing ability. ES/MS-MS is ideally suited to achieve unequivocal identification of high-affinity ligands by this method.

The LAS-MS system was calibrated with DDEY(mY)-LIP, a peptide with high affinity for PTP-1B as determined by the iterative approach. GST-PTP-1B was immobilized to glutathione-Agarose and incubated with DDEY(mY)LIP. Mass spectrometry was performed on three successively more stringent elution steps: (1) unbound, (2) high-salt wash, and (3) low-pH wash. The standard peptide was tightly bound and eluted in the low pH wash at the expected molecular ion [$(M + H)^+$] of *m*/*z* 1129.2 (data not shown). ES-MS/MS analysis of this molecular ion confirmed its sequence.

For the affinity selection process, the most effective inhibitory O_{-3} sublibrary, DDXX(mY)LIP, was chosen for further study. ES-MS analysis of the intact library revealed the diversity of the 400 components of the library, and the peak distribution agreed well with theoretical calculations (Figure 2A). When PTP-1B was screened with this library, most of the library components bound weakly, eluting in the unbound and high-salt washes (not shown). The low-pH acid elution step, however, contained a subset of the library which was affinity selected (Figure 2B). A "relative selection" was used to rank order these library components (Table 1), listed as the molecular ions found. The peak which had the greatest relative selection eluted with a molecular ion [$(M + H)^+$] of *m*/*z* 1129.2, which was a peak corresponding to the amino acid combination E/Y. Confirming this assignment, the ES-MS/MS spectrum of this peak gave a fragmentation pattern that corresponded to the sequence DDEY(mY)LIP (Figure 3). Strikingly, the peaks which had the highest relative selection—EY, DY, and EF—corresponded to peptides containing aromatic residues in the −1 subsite and acidic residues in the −2 subsite, in good agreement with the results obtained from the iterative approach. The information from the ES-MS/MS technique clearly established an unequivocal identification of the sequence of the high-affinity peptides.

Peptides containing tryptophan at the −1 subsite were not seen by the LAS-MS method, but were found to be strong inhibitors of PTP-1B catalysis when tested as individual compounds (round three of the iterative approach). Since it is well documented that the indole portion of the tryptophan side chain is susceptible to oxidation and destruction, this destruction would be magnified within the context of a mixture of 400 peptides and could account for its absence in the library screening. To confirm this, we analyzed the ES-MS spectrum of the O_{-3} sublibrary DDXX(mY)LIP (Figure 2A) for the presence of unambiguous peaks containing a tryptophan residue. As predicted, this spectrum showed a lack of these peptides at concentrations sufficient for ES-MS detection. In contrast, MS performed on authentic DDEW(mY)LIP indicated a normal detection sensitivity for this peptide, ruling out the possibility of a MS detection problem of tryptophan-containing peptides.

Table 1 [a]

$X_{-2}X_{-1}$	$(M + H)^+$	relative selection[b] (%)	ES-MS/MS
EY	1129.3	9	yes
DY	1115.3	5	yes
NY	1114.3	5	yes
EF	1113.3	1	yes
(I,L,hP)[c]Y	1113.3	1	yes
D/F,[d] V/Y,M/M	1099.3	1	no
A/Y,S/F,H/P	1071.3	0.8	no
D/(I,L,hP),E/V,N/N,M/P	1065.3	0.6	no
A/F,S/M	1055.1	0.5	no
E/(I,L,hP),N/(Q/K)	1079.3	0.3	no
G/(I,L,hP),V/A	1007.3	0.2	no
A/(I,L,hP),S/P	1021.2	0.05	no
T/(I/L/hP),V/D	1051.3	0.05	no

[a] The peptides derived from the experiment of Figure 3 were enumerated by ES-MS and, as noted in the table, by ES-MS/MS. ES-MS alone does not resolve sequences (i.e., AY vs. YA) and in these cases, unresolved sequences are listed with a slash as A/Y. Likewise, ES-MS does not distinguish sequences of identical mass (AF and SM) and this potential redundancy is represented as lists separated by comas. Finally, building blocks of identical mol. wt. (I, L, hP) are listed in series in parentheses. The LAS-MS experiment was performed on three occasions with similar results and similar analysis have been performed with three other sublibraries (X_{-3} = G, V, and E), and support the same conclusions. [b] MS is not considered a quantitative analytical tool, for example due to variations in the response factor in the electrospray MS detection. Furthermore, where there is a redundancy of library components at a particular molecular weight, assessment of one of these components is difficult to estimate due to the redundancy. Nevertheless, to estimate selectivity of the components isolated by affinity selection, a "relative selection" factor is presented. It is the normalized area of a particular component found in the pH 3.5 elution divided by the area of that component in the original library, without considering potential redundancy. A nonredundant component, therefore, would have a theoretical limit of 100% and any component represented by a redundancy in the original library would have a theoretical maximum less than 100%, as dictated by its relative contribution to the library peak signal. [c] Amino acids with identical masses (i.e., L = I = hP). [d] Unresolved sequence redundancies.

Phosphotyrosine Phosphatase Analysis. An interesting question arose as to whether the optimized inhibitor would correspond to the optimized substrate as well. To investigate the contributions to catalytic activity by individual subsites, "alanine-scanning" of the substrate peptide $D_{-4}A_{-3}D_{-2}E_{-1}pYL_{+1}I_{+2}P_{+3}$ ($EGFR_{988-995}$) was performed with GST-PTP-1B. This peptide was chosen as it is a good substrate for PTP-1B and because a substantial amount of work of this type has been performed with this substrate using the rat and a bacterial enzyme (*17*). In addition, a derivative containing an aspartic acid substitution at the −3 subsite was prepared, and the release of inorganic phosphate by PTP-1B was determined using the Malachite Green enzyme assay (Figure 4). Substitutions which had a significant effect on the specific activity of PTP-1B for substrate (>50%) included subsites −2 > −1 > −3, with alanine being preferred in the −1/−2 subsites and aspartic acid being preferred in the −3 subsite. Substitutions which had only a modest effect on the specific activity of PTP-1B for substrate (<50%) included subsites +2 > −4 > +1 > +3, with alanine being preferred in the +1/+2 subsites, aspartic acid being preferred in the −4 subsite, and proline being preferred in the +3

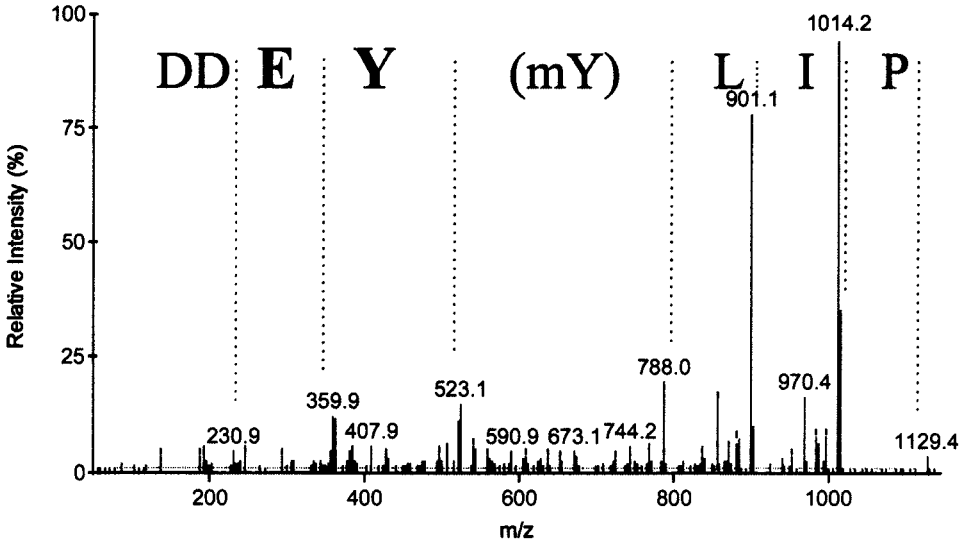

FIGURE 3: ES-MS/MS spectrum of DDEY(mY)LIP after library affinity selection. Library O_{-3}; $DD_{-3}X_{-2}X_{-1}$(mY)LIP (200 μM) was evaluated by LAS-MS as described in the Materials and Methods. The molecular ion [$(M + H)^+$] of *m/z* 1129.2 from the affinity elution was monitored by ES-MS/MS with a scan range of 50 (parent ion + ca. 10) *m/z* units. The b-ions relate to ions formed upon cleavage at the peptide bond. The amino acid sequence is designated above the b-ions in the mass spectrum.

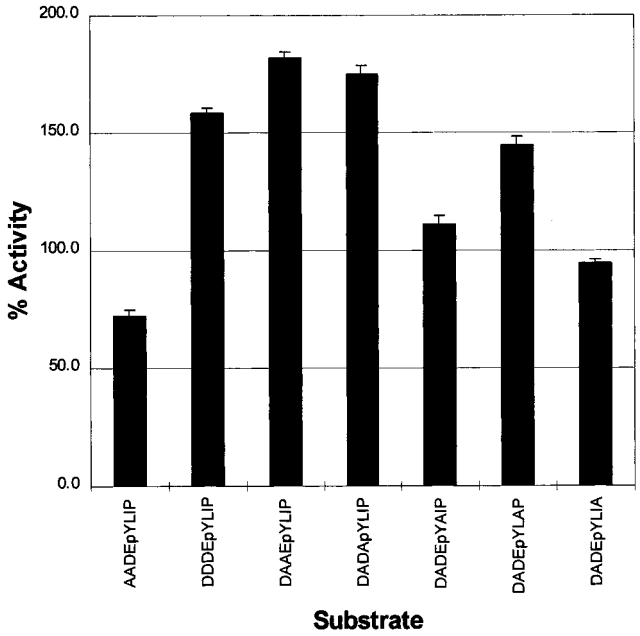

FIGURE 4: Analysis of PTP-1B activity versus various phosphopeptides. The activity of PTP-1B with different substrate variations based upon the peptide DADEpYLIP are presented. Reaction mixtures contained 20 nM GST-PTP-1B and 200 μM phosphopeptide. Specific activity was determined using the Malachite Green Assay (Materials and Methods) and is expressed relative to DADEpYLIP. The results are the average (±SEM) of three independent experiments each performed in triplicate.

subsite. Thus, in agreement with previous results using related PTPases, substitutions at subsites proximal and N-terminal to the phosphorylation site appear to have a greater effect on catalytic activity than those proximal and C-terminal. However, based upon this information, the potential importance of subsites proximal and C-terminal, particularly the +2 subsite, should not be overlooked.

Structural Model. The present information indicating a preference for aromatic residues in the −1 position (Phe, Trp, and Tyr) was considered in light of the crystal structure of human PTP-1B. This structure was established as an inactive enzyme (active-site-mutated Cys215Ser) with a bound hexapeptide substrate, originally presented by Jia et al. (*33*). As shown in Figure 5, Arg 47 is a key contact residue for the −1 position (glutamic acid of the substrate). Assuming the inhibitor peptides elucidated in this study bind in a similar fashion as this substrate, a charged residue—Arg 47—is in close proximity to an aromatic ring of the inhibitor. A stacking configuration may serve to stabilize this interaction. An alternative explanation would involve a conformational change of the enzyme or of the inhibitor, relative to the substrate represented in the model, to take advantage of the proximity of aromatic residue Tyr 46.

DISCUSSION

We sought to elucidate the subsite preferences of human PTP-1B by screening a synthetic combinatorial library of ~8000 components using nonhydrolyzable peptidomimetics of PTP-1B substrates. Using an iterative approach, DDEW-(mY)LIP was found to be the most effective inhibitory sequence. Similar sequences with aromatic −1 and acidic −2 residues served as potent inhibitors. Consistent with these observations, when a LAS-MS approach was used, DDEY(mY)LIP, DDDY(mY)LIP, and DDEF(mY)LIP were preferentially bound. Thus, two cross-validating screening approaches gave similar results and demonstrated a novel preference of PTP-1B for aromatic residues at the −1 subsite. Interestingly, Puius et al. (*34*) have recently documented a bi-aromatic PTPase inhibitor. The selection of acidic residues in the −2 subsite agreed well with previous studies (*17*) and with the PTP-1B/$EGFR_{988-993}$ cocrystal structure (*33*).

While the focus of this work was on inhibitor design, it was of interest to attempt to ask whether the activity of PTP-1B toward pY-containing substrates would correlate with the subsite preferences for inhibitors containing mY as a pY mimic. To do so, we compared the results obtained from the phosphatase assay with those from the combinatorial screening (iterative approach). Both approaches showed a clear preference for aspartic acid over alanine at the −3

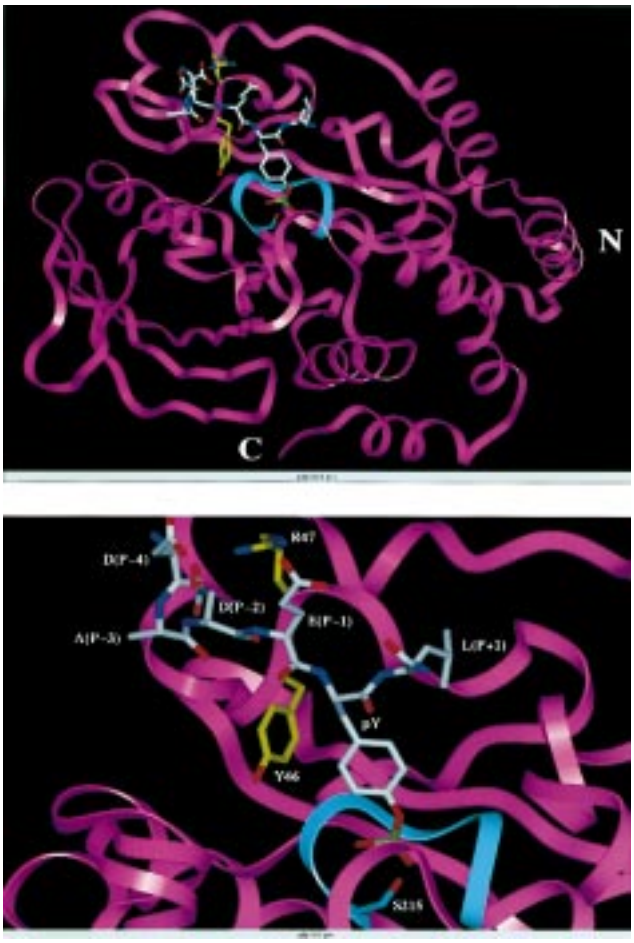

FIGURE 5: A structural model of human PTP-1B. The model was constructed from the crystal structure of human PTP-1B (Cys215Ser) with bound phosphopeptide substrate, as reported by Jia et al. (*33*). The structure was generated using the Insight II program (MSI, Burlington, MA). The main-chain backbone of PTP-1B is represented as a ribbon diagram in lavender. The binding motif region from Cys 215 to Arg 221 is colored in blue. The hexapeptide substrate, DADEpYL-amide is represented as a render molecule (gray, blue and red). The side chain of two essential residues contacting the −1 subsite, Arg 47 and Tyr 46, are highlighted as stick representations (colored yellow). The top picture is a view encompassing the whole enzyme and the lower picture shows an enlargement of the −1 subsite.

subsite, but discrepancies arose at the −1 and −2 subsites. When phosphatase activity was measured using pY-containing peptides, a clear preference for alanine was shown at subsites −1 and −2, whereas, when inhibition activity was measured using mY-containing peptides, glutamic acid was preferred over alanine at subsite −1 and aspartic acid was preferred over alanine at subsite −2. Differing synergistic effects with neighboring residues within the two approaches could account for these discrepancies and may attest to the powers of the combinatorial approach. Alternatively, generalizations from the principles which govern inhibitor binding to the principles which govern substrate specificity may not be valid. Although attractive, there is little reason to assume that important biological substrate sequences for PTP-1B will correspond to its highest affinity peptidomimetic inhibition. Efficient substrate catalysis is dependent on both affinity and rapid turnover; the two properties need not be directly correlated. Another distinction which was considered is the potential influence of the phosphotyrosine mimic on neighboring residues. Although evidence has been presented that the difluoromethylphosphonate (f2pmp-F) and the malonyltyrosine (mY) mimics are able to correctly position in the phosphotyrosine-binding site (*23*, *24*), only recently have studies compared optimum inhibitor design using different phosphotyrosine mimics. Huyer et al. (*35*) studied the optimization of PTP-1B-binding sequences using the f2pmp-F mimic with an affinity-based method employing gel filtration separation on small peptide mixtures. Although the rank order of the certain building blocks in each subsite is subtly different that those reported here using mY as the mimic, overall the results are quite consistent and imply the nature of the mimic is not playing a large role to influence proximal sites.

A number of observations have shown that the specificity of PTPs is determined by a composite of different molecular interactions, both near the active site and outside the catalytic domain, involving both regulatory domains and localization domains. Examples of localization domains—domains outside the catalytic domain which assist binding to selective substrates—are numerous, for both phosphatases and kinases. In fact, owing to these multiple mechanisms, the initial impression that protein phosphatases were promiscuous for substrates has been shown to be incorrect. Clear examples of substrate specificity have been documented (*36*, *37*). In the same fashion, the interaction domain of the substrate includes more than the residues within three amino acids from the phosphotyrosine group. However, a large binding energy can be attributed to this proximal area, as documented by comparing phosphopeptide substrates of different length (*38*).

Many different combinatorial methods can be used to study substrate preferences, including encoded (*39*−*41*), phage display (*42*−*45*), one-bead/one-peptide (*22*, *46*), termination synthesis (*47*), and spatially addressed (*48*, *49*) approaches. The iterative/LAS-MS approach, as first demonstrated for tyrosine phosphatases in this communication, has certain advantages because it supplies a detailed understanding of amino acid preferences at the subsite level and, in addition, it utilizes library components which are free in solution rather than tethered or tagged. The powers of the iterative approach are to move rapidly to a high-affinity ligand and to rank order the contributions of different subsites. Such information creates a powerful matrix for the design and optimization of isoform-specific enzyme inhibitors. While the results presented here and from numerous other papers attest to the power of the iterative approach, no claim is made that the absolute optimal inhibitor structure has been derived. In this regard, it should be noted that, in the case of PTP-1B or any PTPase, there is little concrete information regarding the importance of the positive subsites. Indeed, in view of the results obtained from alanine scanning, optimization of the +2 subsite may further improve the inhibitor structure.

The iterative approach can be criticized because it optimizes one residue at a time and does not fully take into account the combinatorial effects of interactions between distinct residues. In general, confirmatory methods have not been reported. Isolation of the same sequence by the LAS-MS method from a large library cross-validates the iterative approach and has strong confirmatory implications for the optimization process.

A similar analysis with different members of the PTPase superfamily is currently in progress. Preliminary results

suggest the potential for specifically targeting a single PTPase, due to its unique subsite preference, and attests to the utility of this approach. Furthermore, characterization of phosphopeptides containing the inhibitor subsite preferences presented herein could provide important insights into the specific biological substrates for PTP-1B.

REFERENCES

1. Walton, K. M., Dixon, J. E. (1993) Protein tyrosine phosphatases. *Annu. Rev. Biochem. 62*, 101−20.
2. Barford, D., Jia, Z., and Tonks, N. K. (1995) Protein tyrosine phosphatases take off. *Nat. Struct. Biol. 2* (12), 1043−53.
3. Guan, K. L. (1990) Dixon, J. E. (1990) Protein tyrosine phosphatase activity of an essential virulence determinant in Yersinia. *Science 249* (4968), 553−6.
4. Guan, K. L., Broyles, S. S., and Dixon, J. E. (1991) A Tyr/Ser protein phosphatase encoded by vaccinia virus. *Nature 350* (6316), 359−62.
5. Bolin, I.; Wolf-Watz, H. (1988) The plasmid-encoded Yop2b protein of Yersinia pseudotuberculosis is a virulence determinant regulated by calcium and temperature at the level of transcription. *Mol. Microbiol. 2* (2), 237−45.
6. Charbonneau, H., et al. (1989) Human placenta protein-tyrosine-phosphatase: amino acid sequence and relationship to a family of receptor-like proteins. *Proc. Natl. Acad. Sci. U.S.A. 86* (14), 5252−6.
7. Tonks, N. K., Diltz, C. D., Fischer, E. H. (1988) Characterization of the major protein-tyrosine-phosphatases of human placenta. *J. Biol. Chem. 263* (14), 6731−7.
8. Tonks, N. K., Diltz, C. D., Fischer, E. H. (1988) Purification of the major protein-tyrosine-phosphatases of human placenta. *J. Biol. Chem. 263* (14), 6722−30.
9. Goldstein, B. J. (1996) Protein-Tyrosine Phosphatases and the Regulation of Insulin Action. in *Diabetes Mellitus* (LeRoith, D., Taylor, S. I., Olefsky, J. M., Eds.) p 174−186, Lippencot-Raven.
10. Kenner, K. A., et al. (1996) Protein-tyrosine phosphatase 1B is a negative regulator of insulin- and insulin-like growth factor-I-stimulated signaling. *J. Biol. Chem. 271* (33), 19810−6.
11. Ahmad, F.; Goldstein, B. J. (1995) Alterations in specific protein-tyrosine phosphatases accompany insulin resistance of streptozotocin diabetes. *Am. J. Physiol. 268* (5 Pt. 1), E932−40.
12. Ahmad, F., Considine, R. V., and Goldstein, B. J. (1995) Increased abundance of the receptor-type protein-tyrosine phosphatase LAR accounts for the elevated insulin receptor dephosphorylating activity in adipose tissue of obese human subjects. *J. Clin. Invest. 95* (6), 2806−12.
13. Ahmad, F.; Goldstein, B. J. (1997) Functional association between the insulin receptor and the transmembrane protein-tyrosine phosphatase LAR in intact cells. *J. Biol. Chem. 272* (1), 448−57.
14. Ahmad, F., et al. (1997) Improved sensitivity to insulin in obese subjects following weight loss is accompanied by reduced protein-tyrosine phosphatases in adipose tissue. *Metabolism 46* (10), 1140−5.
15. Roth, R. A. (1988) Insulin and insulin-like growth factor receptors and responses. *Cold Spring Harbor Symp. Quant. Biol. 53* (1), 537−543.
16. Zhang, Z. Y., et al. (1993) Substrate specificity of the protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U.S.A. 90* (10), 4446−50.
17. Zhang, Z. Y.; Dixon, J. E. (1994) Protein tyrosine phosphatases: mechanism of catalysis and substrate specificity. *Adv. Enzymol. Relat. Areas Mol. Biol. 68*, 1−36.
18. Milarski, K. L., et al. (1993) Sequence specificity in recognition of the epidermal growth factor receptor by protein tyrosine phosphatase 1B. *J. Biol. Chem. 268* (31), 23634−9.
19. Hippen, K. L., et al. (1993) Acidic residues are involved in substrate recognition by two soluble protein tyrosine phosphatases, PTP-5 and rrbPTP-1. *Biochemistry 32* (46), 12405−12.
20. Cho, H. J., et al. (1991) Purification and characterization of a soluble catalytic fragment of the human transmembrane leukocyte antigen related (LAR) protein tyrosine phosphatase from an Escherichia coli expression system. *Biochemistry 30* (25), 6210−6.
21. Agostinis, P., et al. (1996) A comparative study of the phosphotyrosyl phosphatase specificity of protein phosphatase type 2A and phosphotyrosyl phosphatase type 1B using phosphopeptides and the phosphoproteins p50/HS1, c-Fgr and Lyn. *Eur. J. Biochem. 236* (2), 548−57.
22. Lam, K. S., et al. (1991) A new type of synthetic peptide library for identifying ligand-binding activity [published errata appear in Nature 1992 Jul 30, 358(6385), 434; and 1992 Dec 24−31, 360(6406), 768]. *Nature 354* (6348), 82−4.
23. Kole, H. K., et al. (1995) Protein-tyrosine phosphatase inhibition by a peptide containing the phosphotyrosyl mimetic, L.-O-malonyltyrosine. *Biochem. Biophys. Res. Commun. 209* (3), 817−22.
24. Ye, B., et al. (1995) L-O-(2-malonyl)tyrosine: a new phosphotyrosyl mimetic for the preparation of Src homology 2 domain inhibitory peptides. *J. Med. Chem. 38* (21), 4270−5.
25. Pinilla, C., Appel, J. R., and Houghten, R. A. (1993) Synthetic peptide combinatorial libraries (SPCLs): identification of the antigenic determinant of beta-endorphin recognized by monoclonal antibody 3E7. *Gene 128* (1), 71−6.
26. Ostresh, J. M., et al. (1996) Generation and use of nonsupport-bound peptide and peptidomimetic combinatorial libraries. *Methods Enzymol. 267*, 220−34.
27. Houghten, R. A., et al. (1991) Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature 354* (6348), 84−6.
28. Kelly, M. A., et al. (1996) Characterization of SH2-ligand interactions via library affinity selection with mass spectrometric detection. *Biochemistry 35* (36), 11747−55.
29. Van Veldhoven, P. P., and Mannaerts, G. P. (1987) Inorganic and organic phosphate measurements in the nanomolar range. *Anal. Biochem. 161* (1), 45−8.
30. Lanzetta, P. A., et al. (1979) An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem. 100* (1), 95−7.
31. Barford, D., et al. (1994) Purification and crystallization of the catalytic domain of human protein tyrosine phosphatase 1B expressed in Escherichia coli. *J. Mol. Biol. 239* (5), 726−30.
32. Zhang, W. R., et al. (1994) Molecular cloning and expression of a unique receptor-like protein-tyrosine-phosphatase in the leucocyte-common-antigen-related phosphate family. *Biochem. J. 302* (Pt 1), 39−47.
33. Jia, Z., et al. (1995) Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science 268* (5218), 1754−8.
34. Puius, Y. A., et al. (1997) Identification of a second aryl phosphate-binding site in protein- tyrosine phosphatase 1B: a paradigm for inhibitor design. *Proc. Natl. Acad. Sci. U.S.A. 94* (25), 13420−5.
35. Huyer, G. K. J., Moffat, J., Xamboni, R., Jia, Z., Gresser, M. J., and Ramachandran, C. (1998) Affinity Selection from Peptide Libraries to Determine Substrate Specificity of Protein Tyrosine Phosphatases. *Anal. Biochem. 258*, 19−30.
36. Flint, A. J., et al. (1997) Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U.S.A. 94* (5), 1680−5.
37. Garton, A. J., and Tonks, N. K. (1994) PTP−PEST: a protein tyrosine phosphatase regulated by serine phosphorylation. *EMBO J 13* (16), 3763−71.
38. Zhang, Z. Y., et al. (1994) Protein tyrosine phosphatase substrate specificity: size and phosphotyrosine positioning requirements in peptide substrates. *Biochemistry 33* (8), 2285−90.

39. Needels, M. C., et al. (1993) Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc. Natl. Acad. Sci. U.S.A. 90* (22), 10700−4.
40. Ohlmeyer, M. H., et al. (1993) Complex synthetic chemical libraries indexed with molecular tags. *Proc. Natl. Acad. Sci. U.S.A. 90* (23), 10922−6.
41. Brenner, S., and Lerner, R. A. (1992) Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. U.S.A. 89* (12), 5381−3.
42. Cull, M. G., Miller, J. F., and Schatz, P. J. (1992) Screening for receptor ligands using large libraries of peptides linked to the C terminus of the lac repressor. *Proc. Natl. Acad. Sci. U.S.A. 89* (5), 1865−9.
43. Cwirla, S. E., et al. (1990) Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. U.S.A. 87* (16), 6378−82.
44. Goodson, R. J., et al. (1994) High-affinity urokinase receptor antagonists identified with bacteriophage peptide display. *Proc. Natl. Acad. Sci. U.S.A. 91* (15), 7129−33.
45. Smith, G. P. (1995) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science 228* (4705), 1315−7.
46. Lebl, M., et al. (1995) One-bead-one-structure combinatorial libraries. *Biopolymers 37* (3), 177−98.
47. Youngquist, R. S., et al. (1994) Matrix-assisted laser desorption ionization for rapid determination of the sequences of biologically active peptides isolated from support- bound combinatorial peptide libraries. *Rapid Commun. Mass Spectrom. 8* (1), 77−81.
48. Fodor, S. P., et al. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science 251* (4995), 767−73.
49. Gallop, M. A., et al. (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem. 37* (9), 1233−51.

BI981427+