

Substrate specificity of the protein tyrosine phosphatases

(enzyme kinetics)

ZHONG-YIN ZHANG*, ANDREA M. THIEME-SEFLER†, DEREK MACLEAN†, DENNIS J. MCNAMARA†, ELLEN M. DOBRUSIN†, TOMI K. SAWYER†, AND JACK E. DIXON*‡

*Department of Biological Chemistry, Medical School, Walther Cancer Institute, The University of Michigan, Ann Arbor, MI 48109; and †Department of Chemistry, Parke-Davis Research Division, Warner-Lambert Company, Ann Arbor, MI 48105

Communicated by M. J. Coon, February 18, 1993 (received for review December 23, 1992)

ABSTRACT The substrate specificity of a recombinant protein tyrosine phosphatase (PTPase) was probed using synthetic phosphotyrosine-containing peptides corresponding to several of the autophosphorylation sites in epidermal growth factor receptor (EGFR). The peptide corresponding to the autophosphorylation site, EGFR_{988–998}, was chosen for further study due to its favorable kinetic constants. The contribution of individual amino acid side chains to the binding and catalysis was ascertained utilizing a strategy in which each amino acid within the undecapeptide EGFR_{988–998} (DADEpYLIPQQG) was sequentially substituted by an Ala residue (Ala-scan). The resulting effects due to singular Ala substitution were assessed by kinetic analysis with two widely divergent homogeneous PTPases. A “consensus sequence” for PTPase recognition may be suggested from the Ala-scan data as DADEpYAAPA, and the presence of acidic residues proximate to the NH₂-terminal side of phosphorylation is critical for high-affinity binding and catalysis. The K_m value for EGFR_{988–998} decreased as the pH increased, suggesting that phosphate dianion is favored for substrate binding. The results demonstrate that chemical features in the primary structure surrounding the dephosphorylation site contribute to PTPase substrate specificity.

The tyrosine phosphorylation “status” of a cell is maintained by protein tyrosine kinases (PTKs) and protein tyrosine phosphatases (PTPases) (1, 2). The first PTPase to be purified and sequenced was a 35-kDa protein (PTP 1B) from human placenta (3, 4). PTP1B was shown to share amino acid sequence homology with the cytoplasmic domain of the leukocyte cell surface glycoprotein CD45 (5). CD45 was subsequently shown to have tyrosine phosphatase activity (6). This discovery drew attention to the possible biological role of the PTPases in controlling signal transduction within the cell. Isolation and characterization of cDNA by low stringency hybridization and polymerase chain reaction demonstrated that PTPases constitute a large diversified family of catalysts that can be divided into two structurally distinct groups. One group generally has an extracellular domain, transmembrane spanning region, as well as two duplicated cytoplasmic PTPase domains. The other group of PTPases corresponds to the intracellular family of enzymes that have a single PTPase domain (2).

Many PTPases have been cloned but limited information is available about their functions, mechanisms of catalysis, or substrate specificities. One of the central questions in protein phosphorylation is how kinases and phosphatases distinguish the diversity of substrates that they encounter in the cell. In the case of protein kinases, the amino acid sequence surrounding the phosphorylation site plays a crucial role in determining substrate specificity (7, 8). For example, the cAMP-dependent protein kinase has a strong preference for

phosphorylation of Ser residues that are located two or three residues to the COOH-terminal side of basic amino acids (most commonly Arg). Tyrosine kinases phosphorylate tyrosine residues that are preceded by several acidic amino acid residues (9, 10).

Detailed studies on the substrate specificity of PTPases have been hampered by the lack of highly purified enzymes and stoichiometrically phosphorylated substrates. Therefore, artificial substrates such as *p*-nitrophenylphosphate, phosphorylated lysozyme, or myelin basic protein have been used to characterize the PTPases (11). Recent advances in solid-phase peptide synthesis have resulted in the successful preparation of stoichiometrically tyrosine-phosphorylated peptides (12, 13). In general, two types of assay have been employed for the synthetic phosphopeptides. One is to follow the PTPase-catalyzed hydrolysis by measuring the production of inorganic phosphate using the malachite green colorimetric assay (14), and the other is to follow the breakdown of peptide substrates by HPLC, since the phosphorylated and dephosphorylated forms have different retention times (15). These two techniques are discontinuous assays and the procedures are laborious.

We have recently developed expression and purification schemes to obtain large amounts of homogeneous PTPases (16, 17). We have also reported a sensitive continuous assay for PTPases that utilizes phosphotyrosine-containing peptide substrates (18). This assay is based on the differences in the spectrum of the peptide before and after the removal of the phosphate group. The increase in the absorbance at 282 nm or the increase in fluorescence at 305 nm of the peptide upon the action of PTPase can be followed continuously and the resulting progress curve (time course) can be analyzed directly using the integrated form of the Michaelis-Menten equation. This procedure is convenient and efficient, since k_{cat} and K_m values can be obtained in a single run. To analyze the substrate specificity of the PTPases in greater detail, we have utilized peptides corresponding to the autophosphorylation sites of epidermal growth factor receptor (EGFR) to address some of the structural requirements for optimum substrate binding and catalysis. A strategy was employed in which each amino acid within the phosphotyrosine-containing peptide substrate EGFR_{988–998} (DADEpYLIPQQG) (19) was sequentially substituted by Ala (Ala-scan). The resulting effects of singular Ala substitution were evaluated kinetically. In this way, specific contributions to the binding and catalysis by individual residues were ascertained. Two PTPases were employed in these studies, a recombinant bacterial PTPase from *Yersinia* and a recombinant mammalian PTPase originally cloned from rat brain. We report kinetic parameters with the *Yersinia* PTPase and the recombinant mammalian PTPase and demonstrate that chemical features in the primary struc-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PTPase, protein tyrosine phosphatase; PTK, protein tyrosine kinase; EGFR, epidermal growth factor receptor; SH2, src-homology 2.

‡To whom reprint requests should be addressed.

ture proximate to the phosphorylated Tyr contribute to PTPase substrate specificity.

MATERIALS AND METHODS

Enzyme Preparation. Homogeneous recombinant *Yersinia* PTPase (16) and the mammalian PTPase, PTP1U323 (17), were purified as described. Enzyme and peptide concentrations were determined by amino acid analysis.

Peptide Substrate Preparation. Phosphotyrosine-containing peptides were synthesized, purified, and characterized as described (18). EGFR₁₁₆₇₋₁₁₇₅ (TAENAEpYLRV) was obtained from the Biomedical Research Core Facilities of the University of Michigan.

Enzyme Assay and Data Analysis. All enzyme assays were performed at 30°C in pH-buffered solution with a constant ionic strength (μ) of 0.15 M. All of the peptide substrates were analyzed by the continuous assay (18). For standard spectrophotometric assays, a microcuvette was used. The reaction solution (total volume of 500 μ l) containing an appropriate amount of peptide substrate was incubated at 30°C for at least 15 min before the reaction was started by introducing a catalytic amount of PTPase into the reaction mixture. The entire time course of the PTPase-catalyzed hydrolysis of tyrosine-phosphorylated peptide substrate was recorded by monitoring the increase in absorbance at 282 nm and the Michaelis–Menten kinetic parameters k_{cat} and K_m were determined by analyzing the experimental data through a nonlinear least-squares fit algorithm (20) using the integrated Michaelis–Menten equation:

$$t = p/k_{cat}E_0 + (K_m/k_{cat}E_0)\ln[p_{\infty}/(p_{\infty} - p)],$$

where k_{cat} is the catalytic turnover number, K_m is the Michaelis constant, E_0 is the enzyme concentration, and p and p_{∞} are the product concentration at time t and infinity, respectively. This relationship can then be used directly to analyze an array of experimental t - s data pairs by nonlinear least-squares methods where the parameters k_{cat} and K_m are optimized through minimizing $\sum(t_{exp} - t_{calc})^2$.

RESULTS AND DISCUSSION

Primary Structure Contributes to PTPase Substrate Specificity. Fig. 1 shows a typical time course of the *Yersinia* PTPase-catalyzed hydrolysis of EGFR₉₈₈₋₉₉₈ at pH 6.6, 30°C, monitored by the increase in absorbance at 282 nm. Data from the progress curve were collected and fitted to the integrated form of the Michaelis–Menten rate equation. In most cases, two or more substrate concentrations, at >2 – 5 times the K_m value, were used for each peptide to ensure the validity of the analysis. k_{cat} and K_m values can be obtained simultaneously from a single run. Table 1 summarizes the Michaelis–Menten kinetic parameters of four EGFR peptides with the homogeneous recombinant *Yersinia* PTPase. For comparison, results with six peptides corresponding to the autophosphorylation sites in *neu*, *lck*, and *src* (18) are also included. Although k_{cat} values for all peptide substrates are similar, with the exception of Neu₅₄₆₋₅₅₆, the K_m values vary considerably (*ca.* 30-fold). It is apparent that, as with the protein kinases, the substrate specificity of the PTPases may also be determined, at least in part, by the nature of the residues surrounding the target phosphotyrosine. Due to the differences in sequence and length of the peptide substrates examined, one can not draw any definitive conclusions regarding the structural requirements for substrate recognition. Systematic alteration of amino acids adjacent to the phosphotyrosine in the peptide substrates will shed light on potential recognition elements. We chose the best substrate,

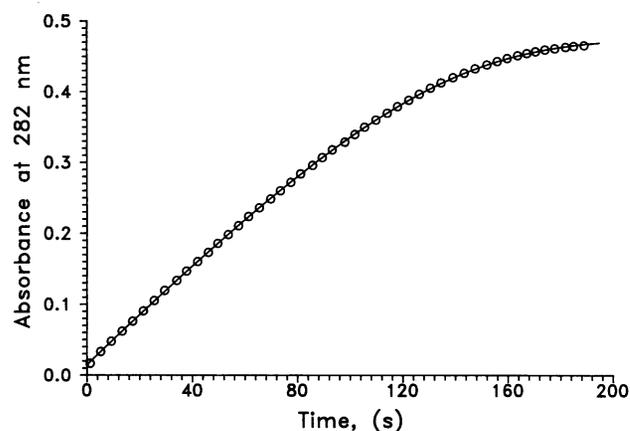


FIG. 1. Progress curve analysis of *Yersinia* PTPase-catalyzed hydrolysis of phosphotyrosine-containing peptide EGFR₉₈₈₋₉₉₈. The hydrolysis reaction was followed by the increase in absorbance at 282 nm at pH 6.6, 50 mM 3,3-dimethylglutarate, 1 mM EDTA, $\mu = 0.15$ M buffer, and 30°C. The peptide substrate concentration was 450 μ M and the enzyme concentration was 3.25 nM. The theoretical curve (solid line) was obtained through a nonlinear least-squares fit algorithm to the experimental data (○) using the integral Michaelis–Menten equation.

EGFR₉₈₈₋₉₉₈, as a template to ascertain specific contributions from individual residues to the binding and catalysis.

Substrate Specificity of the *Yersinia* PTPase. Table 2 summarizes the Michaelis–Menten kinetic parameters of 10 peptides corresponding to EGFR₉₈₈₋₉₉₈. The contribution of specific amino acid side chains to binding and catalysis were determined by systematically replacing each amino acid in the EGFR₉₈₈₋₉₉₈ sequence with Ala (Ala-scan). Kinetic constants were initially obtained with the homogeneous recombinant *Yersinia* PTPase. To simplify the discussion, a nomenclature similar to that used by Kennelly and Krebs (7) will be adopted—i.e., the phosphoacceptor Tyr residue is designated to be at the zero position in the peptide sequence and the adjacent NH₂-terminal and COOH-terminal amino acids will be designated by the numbers NH₂- . . . , -3, -2, -1, 0, +1, +2, +3, . . . -COOH, etc. As shown in Table 2, k_{cat} values of the Ala-scan series of EGFR₉₈₈₋₉₉₈ are relatively constant with the exception of EGFR₉₈₈₋₉₉₈^{E991A} (replacement of E991 with Ala), whereas K_m values vary markedly. The k_{cat}/K_m ratio for EGFR₉₈₈₋₉₉₈ of 2.23×10^7 M⁻¹s⁻¹ at 30°C and pH 6.6 reflects a near diffusion-controlled efficiency for the *Yersinia* PTPase. Substitution of Asp-988 to Ala and Asp-990 to Ala each modestly increased the K_m (2-fold and 5-fold relative to the wild-type peptide, respectively). A more striking effect is observed when Glu-991, which is immediately adjacent to the target tyrosine phosphate, is replaced by Ala (i.e., position -1). This single substitution increases the K_m of the peptide by 63-fold while at the same time reducing its k_{cat} by 2-fold, thus causing a drop of 126-fold in terms of k_{cat}/K_m value (a kinetic parameter of substrate specificity) compared to the wild-type peptide. The importance of the acidic residues as positive regulation elements increases as a function of their proximity to the phosphotyrosine. On the COOH-terminal side of the tyrosine phosphate, substitution of Leu-993 or Gln-996 by Ala has a very modest effect on binding and catalysis, increasing the K_m 1.5- and 2-fold, respectively. Interestingly, the Pro-995 to Ala change increases the K_m by 4-fold (position +3). Finally, replacement of Ile-994, Gln-997, or Gly-998 by Ala has minimal effect on their kinetic behavior. Thus, it is apparent from the Ala-scan that an acidic residue at -1 position is critical for the *Yersinia* PTPase catalysis since its replacement by Ala decreases the substrate specificity by 126-fold. Other important positions are -2 and +3, which are Asp and Pro, respectively. These

Table 1. Kinetic constants for the hydrolysis of phosphorylated peptides by *Yersinia* PTPase at pH 6.6, 30°C

Substrate		K_m , μM	k_{cat} , s^{-1}	$10^{-7} \times k_{\text{cat}}/K_m$, $\text{M}^{-1}\text{s}^{-1}$
DADEpYLIPQQG	EGFR ₉₈₈₋₉₉₈	59.0 ± 4.8	1314 ± 18	2.23
LPVPEpYINASV*	EGFR ₁₀₆₃₋₁₀₇₃	959 ± 50	1349 ± 58	0.141
AEpYLRVAPQS	EGFR ₁₁₇₁₋₁₁₈₀	1990 ± 180	1155 ± 80	0.0580
TAENAEpYLRV	EGFR ₁₁₆₇₋₁₁₇₆	367 ± 24	1146 ± 26	0.312
DAEEpYLVPPQQG†	Neu ₃₄₇₋₃₅₇	78.4 ± 1.6	1348 ± 8.3	1.72
DNLYpYWDQNSS†	Neu ₅₄₆₋₅₅₆	103 ± 28	446 ± 42	0.433
ENPEpYLGLDVPV†	Neu ₅₇₂₋₅₈₃	392 ± 20	1414 ± 41	0.361
EDNEpYTARE†	p56 ^{lck} ₃₉₀₋₃₉₈	2240 ± 450	1676 ± 240	0.0748
TEPQpYQPGE†	p60 ^{src} ₅₂₃₋₅₃₁	390 ± 6.0	1287 ± 9	0.330
AcRRLIEDAEpYAARG†	(RR-src)	189 ± 24	1178 ± 100	0.623

*This peptide corresponds to EGFR₁₀₆₃₋₁₀₇₃, except that residue 1071 Q is replaced with A.

†Results from Zhang *et al.* (18).

conclusions are supported by the fact that Neu₃₄₇₋₃₅₇ (DAEEpYLVPPQQG), which incorporates similar residues at the critical positions noted above (residues that are *different* from EGFR₉₈₈₋₉₉₈ are underlined), has a similar k_{cat}/K_m value ($1.72 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$) under identical conditions (Table 1).

EGFR₁₀₆₂₋₁₀₇₃ (Table 1) has Glu at its -1 position, but the other two important positions—i.e., -2 (Asp or Glu) and +3 (Pro)—have been changed to Pro and Ala, respectively. The consequence was reflected by a K_m of 959 μM for the *Yersinia* PTPase, which agrees well with what would be predicted from the combination of the two in the Ala-scan. Thus the importance of acidic residues NH₂-terminal to the tyrosine phosphate, as well as the individual side chain contributions to the binding and catalysis, are clearly illustrated from the results described above. The sensitivity of PTPase action to the linear sequence surrounding the tyrosine phosphorylation site is also demonstrated by results from kinetic analyses of other physiological relevant peptide substrates corresponding to the phosphorylation sites present within the neu, p56^{lck}, p60^{src}, EGFR, and insulin receptor (14, 18).

The positioning of phosphotyrosine within the peptide sequence is also important for PTPase action. This conclusion can be illustrated by results from another pair of peptide substrates, EGFR₁₁₇₁₋₁₁₈₀ (AEpYLRVAPQS) and EGFR₁₁₆₇₋₁₁₇₆ (TAENAEpYLRV) (Table 1). By extending the NH₂-terminus of the target tyrosine from -2 to -6, and at the same time shortening the COOH-terminus from +7 to +3, we actually observed a >5-fold increase in the k_{cat}/K_m value. Interestingly, a similar observation has been made by Cho *et al.* (14) on the catalytic domain of LAR. In that case, increasing the hexapeptide EGFR₁₁₇₁₋₁₁₇₆ (AEpYLRV; $k_{\text{cat}} = 69 \text{ s}^{-1}$ and $K_m = 4.1 \text{ mM}$) to the undecapeptide EGFR₁₁₆₇₋₁₁₇₇ (TAENAEpYLRVA; $k_{\text{cat}} = 45 \text{ s}^{-1}$ and $K_m = 0.48 \text{ mM}$)

enhances the catalytic efficacy by 5.5-fold. EGFR₁₁₆₇₋₁₁₇₆ lacks the -2 position acidic side chain and the K_m value of 367 μM is compatible with that of EGFR₉₈₈₋₉₉₈^{D990A}. Therefore, greater than two residues NH₂-terminal and three or more residues COOH-terminal to the target tyrosine residue are required for optimal substrate recognition.

The pH dependence of the *Yersinia* PTPase-catalyzed hydrolysis of EGFR₉₈₈₋₉₉₈ is shown in Table 3. The k_{cat}/K_m vs. pH curve is bell-shaped, with the optimal efficiency at pH 6.6, indicating that the phosphate dianion is the preferred substrate. This is similar to the finding that a -2 charge on phosphotyrosine residue is necessary for high-affinity interaction with basic residues within the src-homology 2 (SH2) domain (21, 22). One of the invariant Arg residues present within all PTPases from bacteria to man is a good candidate for binding of the phosphate dianion [i.e., Arg-216, -228, -409, or -437 in the *Yersinia* PTPase amino acid sequence (23)].

Substrate Specificity of the Mammalian PTPase. The mammalian PTPase (PTP1U323; see *Materials and Methods* for details) corresponds to the catalytic domain of PTP1 (17), which shows amino acid sequence identity to intracellular and receptor-like PTPases. Table 2 summarizes kinetic constants for the recombinant PTPase with the same set of EGFR substrates. The k_{cat} values are similar for all of the peptide substrates, whereas their K_m values vary moderately. PTP1 is also a very efficient catalyst since its k_{cat}/K_m ratio of $2.88 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ against EGFR₉₈₈₋₉₉₈ at 30°C and pH 6.6 is approaching the efficiency limited by diffusion events. The Ala-scan on the EGFR₉₈₈₋₉₉₈ does not produce as dramatic an effect on K_m as was observed with the *Yersinia* PTPase. The largest effect due to the Ala substitution is again at position -1 (Glu-991), which decreases the catalytic efficacy by 6.5-fold. Substitution at position +4 (Gln-996) or +6 (Gly-

Table 2. Kinetic constants for the hydrolysis of EGFR₉₈₈₋₉₉₈ Ala-scan peptides by *Yersinia* PTPase and mammalian PTP1 at pH 6.6, 30°C

EGFR ₉₈₈₋₉₉₈ Ala-scan	<i>Yersinia</i> PTPase				Mammalian PTP1			
	K_m , μM	k_{cat} , s^{-1}	$10^{-7} \times k_{\text{cat}}/K_m$, $\text{M}^{-1}\text{s}^{-1}$		K_m , μM	k_{cat} , s^{-1}	$10^{-7} \times k_{\text{cat}}/K_m$, $\text{M}^{-1}\text{s}^{-1}$	
DADEpYLIPQQG	59.0 ± 4.8	1314 ± 18	2.23		2.63 ± 0.37	75.7 ± 1.0	2.88	
AADEpYLIPQQG	132 ± 7.4	1334 ± 12	1.01		4.94 ± 0.68	74.2 ± 0.77	1.50	
DAAEYLIPQQG	298 ± 9.8	1452 ± 36	0.487		5.65 ± 0.74	65.8 ± 1.2	1.16	
DADApYLIPQQG	3773 ± 314	697 ± 76	0.0185		12.4 ± 3.5	54.8 ± 2.2	0.442	
DADEpYAIIPQQG	91.5 ± 5.2	1492 ± 32	1.63		4.05 ± 0.63	57.8 ± 0.67	1.43	
DADEpYLAPQQG	76.4 ± 2.8	1863 ± 26	2.44		4.39 ± 0.51	73.8 ± 0.90	1.68	
DADEpYLIAQQG	250 ± 14	1424 ± 55	0.570		5.78 ± 0.50	73.4 ± 0.78	1.27	
DADEpYLIPAQQG	124 ± 2.8	1565 ± 14	1.26		3.46 ± 0.35	79.9 ± 2.95	2.31	
DADEpYLIPQAG	70.4 ± 3.4	1264 ± 18	1.80		3.88 ± 0.70	53.4 ± 0.61	1.38	
DADEpYLIPQA	55.5 ± 3.0	1256 ± 24	2.26		2.38 ± 0.31	61.4 ± 1.65	2.58	

The Ala-scan, in which the individual amino acid is replaced sequentially by an Ala residue, is highlighted in bold type.

Table 3. pH dependence of the *Yersinia* PTPase-catalyzed hydrolysis of EGFR₉₈₈₋₉₉₈

pH	k_{cat} , s ⁻¹	K_m , μM	$10^{-7} \times k_{\text{cat}}/K_m$, M ⁻¹ s ⁻¹
5.0	2010 \pm 120	483 \pm 40	0.416
5.5	2200 \pm 51	151 \pm 12	1.46
6.0	1950 \pm 36	101 \pm 4.5	1.93
6.6	1314 \pm 14	59 \pm 4.8	2.23
7.0	975 \pm 19	64 \pm 5.1	1.52
7.6	412 \pm 16	58.2 \pm 4.5	0.789
8.0	205 \pm 6.5	44.4 \pm 2.5	0.462
8.7	58.0 \pm 1.6	27.7 \pm 1.0	0.209

998) does not have any noticeable influence on the kinetic properties. Replacement of the additional amino acids in EGFR₉₈₈₋₉₉₈ individually by Ala decreases their k_{cat}/K_m values by ≈ 2 -fold.

PTP1 displays similar stringent requirements in terms of the positioning of phosphotyrosine within the peptide substrates (data not shown) as was noted for the *Yersinia* PTPase—i.e., greater than two residues NH₂-terminal to the target tyrosine residue and three or more residues COOH-terminal to it are required for PTP1 action. The primary structural features surrounding the target tyrosine dephosphorylation site only show a moderate selectivity when PTP1 is used as a catalyst against the same set of peptides used to monitor the *Yersinia* PTPases activity. In an earlier study, three dodecaphosphopeptides corresponding to insulin receptor autophosphorylation sites at pY1146, pY1150, and pY1151 were shown to have approximately equal affinity to PTP1B ($K_m = 1.3$ – $2.5 \mu\text{M}$), suggesting that PTP1B shows no distinct preference for the site of dephosphorylation in these peptides (15). PTP1B from placental membranes shows a K_m of $2 \mu\text{M}$ for the src peptide (RRLIEDAepYAARG) (24). The reason that PTP1 does not display a wider range of substrate selectivity against the peptide substrates is unclear.

Perspective and Implication. Aside from important understandings in terms of substrate specificity gained from the kinetic analysis of phosphotyrosine-containing peptide, one other valuable mechanistic implication is apparent. The k_{cat} values for all of the peptide substrates examined are amazingly similar (Tables 1 and 2) for *Yersinia* PTPase and PTP1. This observation is consistent with both enzymes forming a covalent phosphoenzyme intermediate that is on the catalytic pathway and the breakdown of the intermediate is rate-limiting under the described conditions. In fact, a covalent phosphoenzyme intermediate has been trapped using ³²P-labeled substrates for soluble (25) and receptor-like (26) PTPases. The phosphoenzyme intermediate was then demonstrated to be kinetically competent by fast quench techniques (27) and shown to have a thiophosphate linkage from its chemical characteristics (25) as well as its ³¹P NMR chemical shift (27).

Analysis of the kinetic parameters, k_{cat} , K_m , and k_{cat}/K_m , for the 20 phosphopeptides with the two PTPases represents a start at understanding the “consensus sequence” that characterizes the substrate recognition for the PTPase. A suggested consensus sequence for the *Yersinia* PTPase derived from the Ala-scan and other peptide analyses is DADEpYAAPA. A similar, but less pronounced, trend is also evident for PTP1. The importance of the acidic residues (E or D) at -1 and -2 has been noted along with the proline (P) residue at position $+3$. Ala residues have been placed at other positions in the consensus amino acid sequence. This is meant to imply that substitution of the naturally occurring amino acid with Ala had little effect on catalysis. Substitution of residues other than Ala at these positions could have an effect. Obviously it will be necessary to examine additional

PTPases to determine if the consensus for the *Yersinia* enzyme applies to the entire family of tyrosine phosphatases.

Consensus sequences for kinases have been used to identify autoinhibitory domains involved in the regulation of a number of protein kinases (7). It is interesting to point out that two major phosphoproteins were dephosphorylated by the *Yersinia* PTPase in infected macrophages (35). A PTPase consensus sequence may be extrapolated to identify physiological substrate(s) or inhibitor(s) as well as provide a rational basis for the design of improved substrates for the PTPases. The identification of a consensus sequence for PTPase may also facilitate the design and development of tight-binding inhibitors. Novel, tight-binding PTPase inhibitors will be of importance in the investigation of the regulation of phosphorylation as well as the “cross-talk” mechanism in signal transduction pathways. There may be relevance to the fact that PTKs prefer to phosphorylate Tyr residues located in the vicinity of acidic residues and PTPases prefer to dephosphorylate phosphotyrosines located at similar sites.

It is important to note that the sequences surrounding the sites of phosphorylation–dephosphorylation are only one parameter governing the regulation of PTPases. Evidence suggests that the subcellular localization of PTPases will play a critical role in controlling their substrate specificity. It is clear that many of the PTPases have, in addition to the catalytic domain, a “localization” or targeting domain that directs these catalysts to specific subcellular locations. By definition, this will limit the potential substrates that the PTPase may encounter. We, and others, have described localization domains that target the PTPase to the endoplasmic reticulum (28, 29), the nucleus (30), as well as potential cytoskeletal locations (31). Although the subcellular localization may play a major role in defining the availability of substrate for specific phosphatases, it is likely that it is not the only factor regulating substrate specificity.

Finally, the presence of proteins with SH2 domains is likely to be important in governing the “substrate specificity” and availability of substrates that the PTPases may encounter. SH2-containing proteins show high-affinity binding with phosphotyrosine-containing proteins (21, 32, 33). It has been shown that a linear sequence surrounding a phosphorylated tyrosine is sufficient for SH2 domain recognition (17, 21, 34). These SH2-containing proteins may have an important role in governing the “lifetime” of the phosphotyrosine-containing proteins. Several investigators have noted that the catalytic efficiency of the PTPases exceeds that of the kinases by as much as one to three orders of magnitude (2, 11). One of the roles of an SH2 domain may be to “lengthen” the lifetime of the phosphotyrosine by selective binding and thus preventing dephosphorylation by PTPases (19). Recently, Saltiel and his colleagues (K. Milarski, G. Zhu, and A. Saltiel, personal communication) quantitatively measured the binding of SH2 domains to the phosphorylated EGFR. Interestingly, SH2 domains from a variety of proteins bind to the receptor with varying affinities. The catalytically inactive mammalian PTPase (C215S mutant) (25) also binds very effectively to the phosphorylated receptor. In addition, the binding of PTP1 (C215S) and SH2 domain in the phosphorylated EGFR are mutually exclusive, indicating that they bind to the same site. The series of peptides used in the Ala-scan are capable of competitively displaying the SH2-containing proteins as well as the catalytically inactive phosphatase from the phosphorylated receptor. Sequencing surrounding the site of phosphorylation, localization, or “positional” information dictated by sequences outside of the active site as well as SH2-containing proteins are all likely to contribute to the regulation of PTPases.

We thank Dr. Alan Saltiel (Parke–Davis Research Division, Warner–Lambert Company) for his support and helpful suggestions. We

also acknowledge Dr. Christopher Walsh for his helpful comments on the manuscript. This work was supported by the National Institutes of Health Grant NIDDKD 18849 and the Walther Cancer Institute.

1. Yarden, Y. & Ullrich, A. (1988) *Annu. Rev. Biochem.* **57**, 443–478.
2. Fischer, E. H., Charbonneau, H. & Tonks, N. K. (1991) *Science* **253**, 401–406.
3. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263**, 6722–6730.
4. Charbonneau, H., Tonks, N. K., Kumar, S., Diltz, C. D., Harrylock, M., Coll, D. E., Krebs, E. G., Fischer, E. H. & Walsh, K. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5252–5256.
5. Charbonneau, H., Tonks, N. K., Walsh, K. A. & Fischer, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7182–7186.
6. Tonks, N. K., Charbonneau, H., Diltz, C. D., Fischer, E. H. & Walsh, K. A. (1988) *Biochemistry* **27**, 8695–8701.
7. Kennelly, P. T. & Krebs, E. G. (1991) *J. Biol. Chem.* **266**, 15555–15558.
8. Kemp, B. E. & Pearson, R. B. (1990) *Trends Biochem. Sci.* **15**, 342–346.
9. Patschinsky, T., Hunter, T., Esch, F. S., Cooper, J. A. & Sefton, B. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 973–977.
10. Geahlen, R. L. & Harrison, M. L. (1990) in *Peptides and Protein Phosphorylation*, ed. Kemp, B. E. (CRC, Boca Raton, FL), pp. 239–253.
11. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263**, 6731–6737.
12. Hudson, D. (1988) *J. Org. Chem.* **53**, 617–624.
13. Kitas, E. A., Perich, J. W., Wade, J. D., John, R. B. & Tregear, G. W. (1989) *Tetrahedron Lett.* **30**, 6229–6232.
14. Cho, H., Ramer, S. E., Itoh, M., Winkler, D. G., Kitas, E., Bannwarth, W., Burn, P., Saito, H. & Walsh, C. T. (1991) *Biochemistry* **30**, 6210–6216.
15. Madden, J. A., Bird, M. I., Man, Y., Raven, T. & Myles, D. D. (1991) *Anal. Biochem.* **199**, 210–215.
16. Zhang, Z.-Y., Clemens, J. C., Schubert, H. L., Stuckey, J. A., Fischer, M. W. F., Hume, D. M., Saper, M. A. & Dixon, J. E. (1992) *J. Biol. Chem.* **267**, 23759–23766.
17. Guan, K. L. & Dixon, J. E. (1991) *Anal. Biochem.* **192**, 262–267.
18. Zhang, Z.-Y., Maclean, D., Thieme-Seffler, A. M., Roeske, R. & Dixon, J. E. (1993) *Anal. Biochem.*, in press.
19. Rotin, D., Margolis, B., Mohammadi, M., Daly, R. J., Daum, G., Li, N., Fischer, E. H., Burgess, W. E., Ullrich, A. & Schlessinger, J. (1992) *EMBO J.* **11**, 559–567.
20. Yamaoka, K., Tanigawara, Y., Nakagawa, T. & Uno, T. (1981) *J. Pharmacobio-Dyn.* **4**, 879–885.
21. Domchek, S. M., Auger, K. R., Chatterjee, S., Burke, T. R., Jr., & Shoelson, S. E. (1992) *Biochemistry* **31**, 9865–9870.
22. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. (1991) *Science* **252**, 668–674.
23. Guan, K.-L. & Dixon, J. E. (1990) *Science* **249**, 553–556.
24. Pallen, C. J., Lai, D. S. Y., Chia, H. P., Boulet, I. & Tong, P. H. (1991) *Biochem. J.* **276**, 315–323.
25. Guan, K. L. & Dixon, J. E. (1991) *J. Biol. Chem.* **266**, 17026–17030.
26. Cho, H., Ramer, S. E., Itoh, M., Kitas, E., Bannwarth, W., Burn, P., Saito, H. & Walsh, C. T. (1992) *Biochemistry* **31**, 133–138.
27. Cho, H., Krishnaraj, R., Kitas, E., Bannwarth, W., Walsh, C. T. & Anderson, K. S. (1992) *J. Am. Chem. Soc.* **114**, 7296–7298.
28. Woodford-Thomas, T. A., Rhodes, J. D. & Dixon, J. E. (1992) *J. Cell Biol.* **117**, 401–404.
29. Frangioni, J. V., Beahm, P. H., Shifrin, V., Jost, C. A. & Neel, B. G. (1992) *Cell* **68**, 545–560.
30. McLaughlin, S. & Dixon, J. E. (1992) *J. Biol. Chem.*, in press.
31. Gu, M.-X., York, J. D., Warshawsky, I. & Majerus, P. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5867–5871.
32. Wood, E. R., McDonald, O. B. & Sahyoun, N. (1992) *J. Biol. Chem.* **267**, 14138–14144.
33. Zhu, G., Decker, S. J. & Saltiel, A. R. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9559–9563.
34. Escobedo, J. A., Kaplan, R. R., Kavanaugh, W. M., Turck, C. W. & Williams, L. T. (1991) *Mol. Cell. Biol.* **11**, 981–990.
35. Bliska, J. B., Clemens, J. C., Dixon, J. E. & Falkow, S. (1992) *J. Exp. Med.* **176**, 1625–1630.