

Identification of New Substrates of the Protein-tyrosine Phosphatase PTP1B by Bayesian Integration of Proteome Evidence^{*[5]}

Received for publication, June 24, 2010, and in revised form, November 30, 2010 Published, JBC Papers in Press, December 1, 2010, DOI 10.1074/jbc.M110.157420

Emanuela Ferrari^{‡1}, Michele Tinti^{‡1}, Stefano Costa[‡], Salvatore Corallino[‡], Aurelio Pio Nardoza[‡], Andrew Chatranyamontri[‡], Arnaud Ceol[‡], Gianni Cesareni^{‡5,2}, and Luisa Castagnoli^{‡3}

From the [‡]Department of Biology, University of Rome Tor Vergata, Via della Ricerca Scientifica, 00144 Rome, Italy and

⁵IRCCS Fondazione Santa Lucia, 00143 Rome, Italy

There is growing evidence that tyrosine phosphatases display an intrinsic enzymatic preference for the sequence context flanking the target phosphotyrosines. On the other hand, substrate selection *in vivo* is decisively guided by the enzyme-substrate connectivity in the protein interaction network. We describe here a system wide strategy to infer physiological substrates of protein-tyrosine phosphatases. Here we integrate, by a Bayesian model, proteome wide evidence about *in vitro* substrate preference, as determined by a novel high-density peptide chip technology, and “closeness” in the protein interaction network. This allows to rank candidate substrates of the human PTP1B phosphatase. Ultimately a variety of *in vitro* and *in vivo* approaches were used to verify the prediction that the tyrosine phosphorylation levels of five high-ranking substrates, PLC- γ 1, Gab1, SHP2, EGFR, and SHP1, are indeed specifically modulated by PTP1B. In addition, we demonstrate that the PTP1B-mediated dephosphorylation of Gab1 negatively affects its EGF-induced association with the phosphatase SHP2. The dissociation of this signaling complex is accompanied by a decrease of ERK MAP kinase phosphorylation and activation.

Protein-tyrosine phosphatases (PTPs)⁴ in concert with tyrosine kinases contribute to the maintenance of regulated levels of tyrosine phosphorylation in multicellular organisms (1, 2). Evidence accumulated over the past decade has now led to the recognition that PTPs play specific, active and even dominant roles in setting the levels of tyrosine phosphorylation in cells thus participating in the regulation of many physiological processes including cell growth, tissue differentiation, and intercellular communication (3–5). Disruption of PTPs activity or its dys-

regulation results in a variety of pathologies such as autoimmunity, diabetes, or cancer (4, 6). Despite this recognized importance, the identification of the mediators of the regulatory functions of this enzyme class has been set back by the inherent difficulties in identifying physiological substrates.

Several reports have indicated that the catalytic domains of tyrosine phosphatases, when probed *in vitro*, display an intrinsic, albeit somewhat weak, preference for phosphorylated tyrosines embedded in specific sequence contexts (5). On the other hand, it is also clear that not all the peptides matching these weak consensi are targeted by the phosphatases *in vivo* and that the enzymatic domains are guided to their functional substrates via a network of interactions contributing to position the enzyme and its targets in physical vicinity (5). The relative importance of these two features, enzymatic specificity and network context, for selection of functional phosphatase targets has not been firmly established. On the other hand, it has been shown that the integration of contextual information in predictive strategies can dramatically improve the ability to infer functional targets of the members of the related tyrosine kinase enzyme family (7).

We report here a strategy, aimed at the identification of *in vivo* phosphatase substrates, by combining orthogonal information in a statistically meaningful framework. This is based on the Bayesian integration of two types of evidence: (i) a broad analysis of substrate preference obtained by probing with the phosphatase domain a large number of naturally occurring phosphopeptides arrayed on a glass chip and (ii) a distance matrix detailing the distance between any pair of proteins in a weighted protein interaction graph. This strategy is general and is outlined in Fig. 1. We demonstrate here its application to the identification of new substrates of the PTP1B phosphatase, the best studied member of the PTP superfamily (8–12). Besides, we propose PTP1B as a regulator of the Gab1 interactome.

EXPERIMENTAL PROCEDURES

Array Organization and Binding Assay—As exemplified in supplemental Fig. S1, each glass slide (chip) is organized as three identical replicated arrays (subarrays); each subarray contains 6057 naturally occurring phosphorylated peptides and several control spots (343) arranged in a grid of 6400 positions. 1604 of these are peptides from human proteins that

* This work was supported by grants from the Italian Association for Cancer Research (AIRC), Telethon, and the European Network of Excellence ENFIN.

⌘ Author's Choice—Final version full access.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental “Materials and Methods,” Tables S1–S7, and Figs. S1–S5.

¹ Both authors contributed equally to this work.

² To whom correspondence may be addressed. E-mail: castagnoli@uniroma2.it.

³ To whom correspondence may be addressed. Fax: 39-06-2023500; E-mail: cesareni@uniroma2.it.

⁴ The abbreviations used are: PTP, protein-tyrosine phosphatase; PSSM, position specific scoring matrix; WIC, weighted interactome distance; ROC, receiver operating characteristic; AUC, area under curve.

PTP1B Substrates in the Ras Pathway

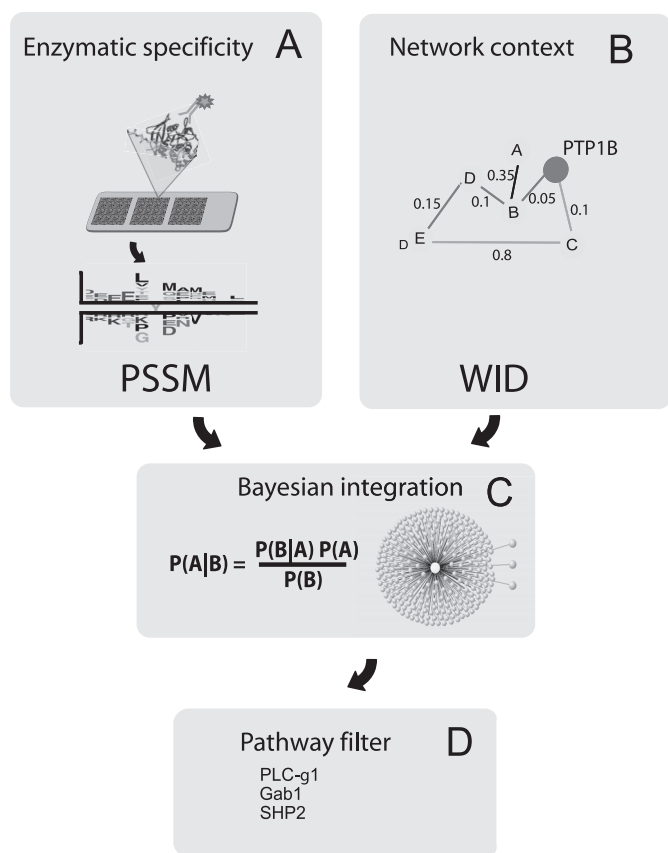


FIGURE 1. Schematic illustration of the strategy used to discover potential PTP1B substrates. *A*, high density peptide chip technology is used to identify the phosphopeptides that have the capacity of binding the trapping mutant catalytic domain of PTP1B. The results are used to create a PSSM to rank the human tyrosine-phosphorylated proteins according to the likelihood of being PTP1B substrates *in vitro*. *B*, a second feature used in the prediction strategy is WID. We first compiled a weighted human interactome by combining the interactions imported from four major protein-protein interaction databases: DIP, HPRD, IntAct, and MINT. For each protein in the human interactome we next calculated WID, a weighted graph distance from PTP1B. *C*, because the two features are largely independent we integrated them by a naive Bayesian approach. The Bayesian framework ranked each phosphorylated human protein with a score reflecting the probability of being a physiological target of PTP1B. *D*, finally we used a Pathway filter to retain only the substrates that had already been described as playing a role in the EGF and insulin pathways.

were described in the literature to be phosphorylated *in vivo*, as curated in the PhosphoELM data base (13). To this list, we added 4453 naturally occurring peptides with a high probability of containing a phosphorylated tyrosine residue according to the NetPhos predictive algorithm (14). The list of phosphopeptides is in [supplemental Table S1](#). Each peptide maintains the same grid position in the three subarrays. The arrays were co-developed with Jerini Peptide Technology within a project supported by the “Interaction Proteome” EU integrated project aimed at the characterization of the recognition specificity of the human SH2 domains.⁵ The glass slides were washed for 1 h at room temperature (RT) in 5 ml of blocking buffer (PBS supplemented with 5% BSA), and incu-

bated with 1 mg/ml of GST-PTP1B trapping mutant for 1 h at RT in 5 ml of blocking buffer. After washing three times with 5 ml of PBS, the glass slides containing the phosphopeptide arrays were incubated for 1 h in the dark with anti-GST Cy5-conjugated antibody (Amersham Biosciences, 1/1000 dilution) in 5 ml of blocking buffer. The chip was extensively washed in 1× PBS and the fluorescence intensity at each array position was measured with a ScanArray Gx Plus instrument (PerkinElmer Life sciences).

Two Sample Logo Visualization—To produce a sequence logo illustrating PTP1B target recognition specificity we used the Two Sample Logo web tool. This software requires as input a positive and a negative set of aligned amino acid sequences. We assembled the positive dataset by selecting all the peptides with a binding intensity higher than the median signal, plus 1 S.D. (300 peptides in total), whereas the peptides with a signal lower by more than 1 S.D. than the median (247) were included in the negative peptide dataset. The same phosphopeptide arrays used to profile the PTP1B specificity were also employed in a recent screening carried out in our laboratory to characterize the interaction network mediated by a representative set of human SH2 domains.⁵ We exploited the results of this larger screening to reduce the risk of including in the negative PTP1B set peptides that did not give a signal in the assay because of technical reasons (*e.g.* poor synthesis or low phosphorylation). To this end we eliminated from the analysis the 194 peptides that had not been bound by any of the SH2 domains tested or by an anti-phosphotyrosine antibody (4G10). The sequence logo of the nonreacting peptides does not show any striking enrichment or bias.

Position Specific Scoring Matrix (PSSM)—To assign to each phosphorylated tyrosine-containing peptide a value reflecting the probability of being dephosphorylated by PTP1B, we have assembled three amino acid frequency matrices: the positive matrix (PM), 20 rows and 13 columns, whose value at row *i* and column *j* is the frequency of amino acid *i* at the *j*th position in the multialignment of the positive peptide set (see previous section). An analogous negative matrix (NM) was obtained from the alignment of 247 “negative peptides”; finally, a total matrix (TM) was compiled by determining the amino acid frequency at each position in the complete list of all the peptides arrayed on the chips ([supplemental Table S3](#)). The score assigned to each query phosphotyrosine containing peptide was then calculated by adding up the overall peptide positions,

$$\sum_j \frac{pm_{ij} - nm_{ij}}{tm_{ij}} \quad (\text{Eq. 1})$$

where pm_{ij} , nm_{ij} , and tm_{ij} are the values for amino acid *i* at position *j* in the positive, negative, and total matrix, respectively (see [supplemental Table S2](#)).

In Vitro Dephosphorylation Assay—T-REx-293 cells were transfected with 10 μg of SRC kinase carrying the Y527F mutation. Cells were collected and lysed in RIPA as described above. The lysate was immunoprecipitated with anti-Gab1, anti-PLCγ-1, and anti-Grb2 antibodies overnight at 4 °C, and then protein A-Sepharose beads were added and incubated for 2 h at 4 °C. After five washings with reaction buffer con-

⁵ M. Tinti, S. Costa, L. Kiemer, M. Miller, F. Sacco, J. Olsen, M. Carducci, S. Paoluzi, C. Workman, N. Blom, K. Machida, C. Thomson, M. Schutkowski, S. Brunak, M. Mann, B. Mayer, L. Castagnoli, and G. Cesareni, manuscript in preparation.

taining 20 mM Hepes, pH 7.4, 150 mM NaCl, and 60 mM EDTA, the dry beads were incubated with 0.3 μ g of PTP1B catalytic domain or 0.3 μ g of GST for 0, 5, 10, and 20 min. The immunoprecipitates were separated by SDS-PAGE, transferred to a nitrocellulose membrane, and immunoblotted with anti-4G10, anti-Gab1, anti-PLC γ -1, and anti-Grb2 antibodies. The catalytic activity of the PTP1B domain was assessed using *p*-nitrophenyl phosphate assay. The reaction was stopped adding 0.1 N NaOH and the enzymatic activity was monitored by measuring the absorbance at 405. The enzyme activity (μ mol/min/ μ g) was calculated applying the Lambert-Beer law.

RESULTS

PTP1B Substrate Specificity—To achieve a comprehensive view of PTP1B target recognition, we used a strategy whereby a domain of interest is challenged with a chip containing the entire collection of peptides that the domain is likely to encounter in the cell. Phosphatases interact only transiently with their substrates, thus making their detection difficult, therefore, to affinity bind the substrates of PTP1B, we use the trapping mutant of the phosphatase, which is a point mutation that maintains specificity but does not allow detachment of the phosphate. To produce the substrates, we made use of a high-throughput technology that we recently developed. This is based on the ability to array on a glass microscope slide (chip) up to 20,000 peptides. For this project, we assembled a chip containing three identical replicas of \sim 6,000 naturally occurring human phosphotyrosine peptides (supplemental Fig. S1). The glass chip was incubated with the PTP1B trapping mutant (12, 15) fused to GST and the interactions were revealed with a secondary fluorescently labeled anti-GST antibody (Fig. 2A). A second chip was probed with GST to identify false-positive peptides that either bind to the tag or to the antibodies used in the assay (results not shown). The signals in the three replica arrays had a Pearson correlation coefficient of 0.89, on average. For each peptide, we used the median of the three replicas as an estimate of its affinity for the trapping mutant. Next, we defined as “binders” (positive set) the peptides whose signal intensity exceeded the median by more than 1 S.D. (positive threshold). Conversely, we defined as “non-binders” the peptides whose signal intensity was lower by more than a standard deviation than the chip median (negative threshold). The 300 binding and the 247 non-binding peptides were used as input for the Two Sample Logo software to visualize the motif characterizing the peptides preferentially recognized by the PTP1B phosphatase (Fig. 2B) (16). Our analysis confirms and extends the findings of several groups (17–19); namely, the sequences of peptides preferentially bound by the catalytic domain of PTP1B are enriched for negatively charged residues at the amino side of the phosphorylated tyrosine. Leucine is the most common residue found at position -1 although residues containing aromatic (Tyr, Phe) or aliphatic (Val, Leu) side chains can also be accommodated at this position. A large fraction of binding peptides have a methionine at $+1$, whereas positions from $+2$ to $+4$ show some, albeit less pronounced, preferences. Most importantly, the large number of peptides tested and the possibility, offered by our assay, to define a negative peptide set,

allow the unprecedented identification of statistically significant amino acid under-representations, at different positions, in the target peptide. The most pronounced signals are the aversion for positively charged side chains at the amino side of the sequence Logo and for Gly and Pro or Pro and negatively charged side chain at the positions immediately preceding or following the phosphotyrosine, respectively.

PTP1B Target Prediction—The phosphopeptide collection arrayed on our chip was originally designed to contain most of the tyrosines known to be phosphorylated *in vivo*. The recent explosion of phosphoproteomic information (20) has made this list incomplete. Thus, we set out to use the experimental information, derived from the peptide array approach, to train a predictive algorithm that could be used to assess the “propensity” to be a PTP1B substrate of any newly discovered phosphotyrosine peptide. Although we also explored an artificial neural network approach, we finally settled on a PSSM method because of its better performance in this instance. The PSSM score (see “Experimental Procedures”) takes into account the residue enrichments in both the positive- and negative-binding peptide datasets. We evaluated the performance of our PSSM-based predictor by three approaches (Fig. 2, C–E). We first assessed the potential of our predictor to reproduce the results of the array experiment by a 10-fold cross-validation approach (test of accuracy) and by measuring the area under the receiver operating characteristics curve (AUC 0.90) (test of specificity), as illustrated in Fig. 2C. As a second independent test, we plotted our quantitative predictions against the PTP1B phosphatase activity recently reported by Barr and co-workers (21) for a panel of 37 phosphopeptides. The Pearson correlation coefficient between the two independent measurements is 0.75 suggesting that the PSSM scores provide a good estimate of the PTP1B activity measured on different peptide sequences, *in vitro* (Fig. 2D).

Phospho-PLC- γ 1, -Gab1, -SHP2, -EGFR, and -SHP1 Are Substrates of PTP1B—Finally, for a limited number of inferred substrates, annotated in the EGF pathway and involved in insulin signaling, we assessed the ability to bind PTP1B-D181A, the substrate trapping mutant of PTP1B, in pulldown assays from extracts of cells stimulated with EGF. We tested seven proteins that are predicted to contain a PTP1B peptide substrate (PLC- γ 1, Gab1, SHP2, EGFR, SHP1, Abl, and Hrs) and two (Grb2 and Endophilin-1) that do not contain any peptide with a high PSSM score. As shown in Fig. 2E, although hardly any phosphorylated protein could be affinity purified with GST alone or with the wild type PTP1B catalytic domain (Fig. 2E), numerous bands were detected in the PTP1B-trapping lane (lane 3). By probing the proteins that were affinity purified with the PTP1B substrate-trapping mutant with specific antibodies, five of seven predicted substrates (PLC- γ 1, Gab1, SHP2, EGFR, and SHP1) were confirmed to specifically bind PTP1B-D181A (Fig. 2E, lane 3). Because we do not know whether in these conditions the Hrs peptide predicted to be a PTP1B substrate is phosphorylated, we cannot conclude whether Hrs is incorrectly inferred to be a PTP1B substrate or rather this negative result is an artifact of the experimental conditions. Grb2 and Endophilin-1, which do not contain phosphotyrosine peptides with a high

PTP1B Substrates in the Ras Pathway

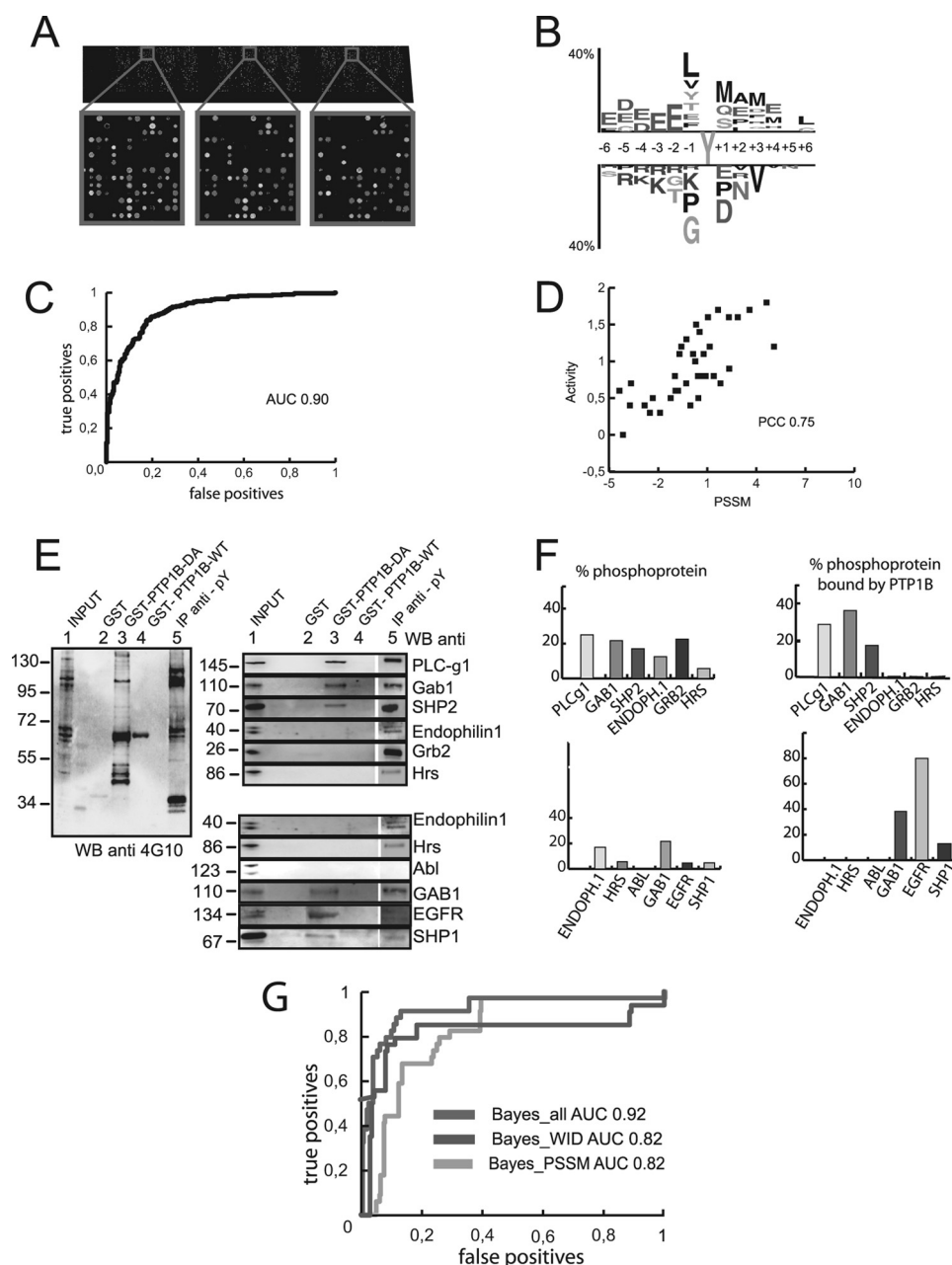


FIGURE 2. PTP1B substrate prediction. *A*, the PTP1B substrate specificity was investigated by probing a high-density phosphopeptide chip with the substrate-trapping mutant of the PTP1B catalytic domain. Corresponding subarray regions ($\sim 2 \times 2$ mm) are enlarged to demonstrate the reproducibility of the signal. *B*, two Sample Logo representation of PTP1B specificity. The logo is divided into two parts. At each position of the *top diagram* the residues (*one letter code*) that are enriched in the "positive peptide set" are represented with a font size relating to their enrichment factor. Conversely, the *bottom diagram* contains the residues that are enriched in the peptide sequences of the negative set. *C*, the performance of the PSSM was evaluated by a 10-fold cross-validation and by plotting a ROC curve, using as positive dataset the 300 binding and as negative dataset the 247 non-binding peptides identified by probing the phosphopeptide chip with the trapping mutant domain of PTP1B. *D*, correlation between PTP1B phosphatase activity and PSSM values. The phosphatase activity of PTP1B when assayed with 37 different peptides (21) was plotted against the peptide PSSM score. *E*, PLC- γ 1, Gab1, and SHP2 bind the substrate-trapping mutant of PTP1B. Bacterially expressed GST, PTP1B trapping mutant domain (GST-PTP1B-DA), or wild type PTP1B catalytic domain (GST-PTP1B-WT), were used in pull-down assays by mixing with 1 mg of lysate from EGF-induced HEK293 cells. The affinity-purified proteins were separated by SDS-PAGE, transferred onto nitrocellulose membranes, and probed with anti-phosphotyrosine antibody 4G10 (*left*) or protein-specific antibodies (*right*). An empty lane in both gel separates the input (*lane 1*) from *lane 2*, containing the GST pull down (negative control). *Lane 3* displays the proteins bound to PTP1B substrate-trapping (the putative PTP1B substrates) and *lane 4* displays the proteins that remain bound to wild type PTP1B. *Lane 5*, in both gels, show proteins immunoprecipitated with the anti-phosphotyrosine antibody 4G10 (to check phosphorylation level). *F*, quantitation of the pull down in *panel E*. The bar graph on the *left* represents the fraction of each protein that is phosphorylated (*i.e.* immunoprecipitated by the anti-phosphotyrosine antibody). The graph on the *right* represents the fraction of phosphorylated protein that is pulled down by the PTP1B substrate trapping mutant. *G*, the curves in the figure compare the ROCs of three differently constructed predictive models: the *blue* and *green* curves refer to models built on a single feature (PSSM or WID, respectively), whereas the *red* curve represents the performance of Bayesian models obtained by integrating both the PSSM and WID scores. The curves shown have been obtained by averaging 10 independently trained classifiers of the same kind by a bootstrapping approach. The AUC distributions of the 10 different models of the three predictors had a standard deviations of 0.02 or less.

PSSM score, are phosphorylated but unable to bind the PTP1B substrate-trapping mutant. Abl, although predicted to be a substrate, is found not to be phosphorylated in our experimental conditions.

Inference of *in Vivo* Substrates—To assess the performance of our PSSM in the inference of *in vivo* PTP1B substrates, we first assembled a list of experimentally verified PTP1B partners by searching the literature for evidence of enzymatic and/or physical interactions between PTP1B and other proteins of the human proteome. By a combination of automatic text mining and manual curation, we have collected information from 84 articles describing 18 PTP1B ligands and 33 PTP1B substrates (see [supplemental Table S4](#)). The set of experimentally verified PTP1B enzymatic substrates was dubbed “positive gold standard.” Due to the intrinsic difficulty of establishing with certainty a set of PTP1B “non-substrate” proteins *in vivo*, the negative golden standard set has been assembled by random sampling (340 proteins) the human phosphotyrosine proteome, under the hypothesis that the majority of human proteins are not PTP1B substrates. The performance of the PSSM in classifying positive and negative examples is reported in Fig. 2G (Bayes_PSSM) as a ROC curve where the percentage of true positive hits is plotted as a function of false positive hits at decreasing PSSM score. The AUC of 0.82 (compare it with 0.90 in Fig. 2C) confirms that features other than peptide substrate specificity must be considered for a correct inference of phosphatase substrates *in vivo*. To this end, we have integrated the *in vitro* evidence encoded in the PSSM, with information about the protein interaction web linking the phosphatase enzyme to its candidate substrates.

Prediction of *in Vivo* PTP1B Substrates by Integration of Recognition Specificity with Network Features—We have first collected in the MINT database all the interactions that can be downloaded from public databases and we have assigned to each interaction a reliability score, ranging from 0 to 1, which takes into consideration all the supporting experimental evidence ([supplemental “Materials and Methods”](#)) (22). The resulting weighted network was dubbed WINT_homo and is available at the MINT download page. Following an analogy between reliability and distance, we considered that proteins connected by high-weight edges can be thought of as being “closer” to each other than proteins connected by low-weight edges. We could thus turn weights into distances (the higher the weight, the lower the distance). Next, we defined the “distance” between two proteins as the length of the shortest weighted path connecting two nodes in the protein interaction graph (Fig. 1B) and we calculated the “weighted interaction distance” (WID) between PTP1B and any other protein present in WINT_homo ([supplemental Table S5](#)). The predictive model trained on this feature has by itself a good performance when used to classify PTP1B substrates and non-substrates (AUC of 0.82, Fig. 2G, Bayes_WID). However, the performance of the Bayesian classifier obtained from the integration of both PSSM, and WID conditional probabilities clearly outperforms (AUC 0.92) the two Bayesian models constructed considering one feature at a time, indicating that each of the features contributes specific information to the

model (Fig. 2G, Bayes_all). We next used the Bayesian classifier to rank all proteins either containing a tyrosine phosphorylation site, as reported in the PhosphoSite and PhosphoELM databases, or connected to PTP1B by at least one path of finite length in the human interactome ([supplemental Table S6](#)). By setting an arbitrary threshold on the Bayesian score, we selected a total of 236 putative phosphatase targets, 17 of which were among the 33 substrates already described in the literature. This corresponds to a sensitivity of 50% at a specificity of 98% (F score = 0.66). Among the proteins predicted to be PTP1B enzymatic targets, we focused on those known to be involved in both EGFR and insulin signaling pathways ([supplemental Table S7](#)). In particular, we decided to further investigate the two predicted targets with the highest score, Gab1 and PLC- γ 1, and another high scoring interactor, SHP2, because of a specific interest of our group in this protein.

Ectopic Expression of PTP1B Specifically Decreases Tyrosine Phosphorylation of Its Inferred Substrates—To determine whether the PLC- γ 1, Gab1, and SHP2 tyrosine phosphorylation levels are altered by PTP1B expression, we ectopically expressed in HEK293 cells the full-length phosphatase (Fig. 3, A and B) and we monitored the phosphorylation levels of the inferred targets upon insulin stimulation. Similar results were obtained after EGF stimulation (results not shown). Aliquots of the cell lysates, transfected with a vector directing the expression of PTP1B or with a control empty vector, were immunoprecipitated with antibodies recognizing PLC- γ 1, Gab1, and SHP2, after separation by gel electrophoresis, the levels of tyrosine phosphorylation were revealed with the anti-4G10 antibody (Fig. 3C). Fig. 3C reports the ratio of the phosphorylated to unphosphorylated forms for each protein analyzed, as a function of time, upon insulin induction. In this experimental setup we have not been able to detect any signal with the 4G10 anti-phosphotyrosine antibody when cell extracts are first immunoprecipitated with an anti-SHP2 antibody, probably as a consequence of the low concentration of phosphorylated SHP2. On the contrary, upon PTP1B overexpression, phosphorylation of PLC- γ 1 and Gab1 sharply decreases when compared with controls where PLC- γ 1 and Gab1 phosphorylations peak at 5 min, after addition of insulin and markedly decline after 20 min.

Likewise, if we first affinity purify the phosphorylated proteins with an anti-phosphotyrosine antibody and then we monitor, by Western blot, the levels of the inferred targets (Fig. 3D), we observe a consistent drop in the phosphorylation of all the three substrates of PTP1B, upon its overexpression. In these conditions, it could be proven also that the SHP2 phosphorylation level is indeed affected by PTP1B. The observed reduction in the phosphorylation levels is specific for the inferred PTP1B targets because, despite PTP1B overexpression, the cell phosphorylation profile does not change appreciably (Fig. 3A) and, in addition, a protein, which received a low score in our predictive algorithm, such as the adapter Grb2, does not show any phosphorylation reduction contingent on PTP1B expression (Fig. 3D). Similar results were obtained when phosphorylation was triggered by the expression of a constitutively active form of the kinase Src, thus supporting the conclusion that the observed reduction in

PTP1B Substrates in the Ras Pathway

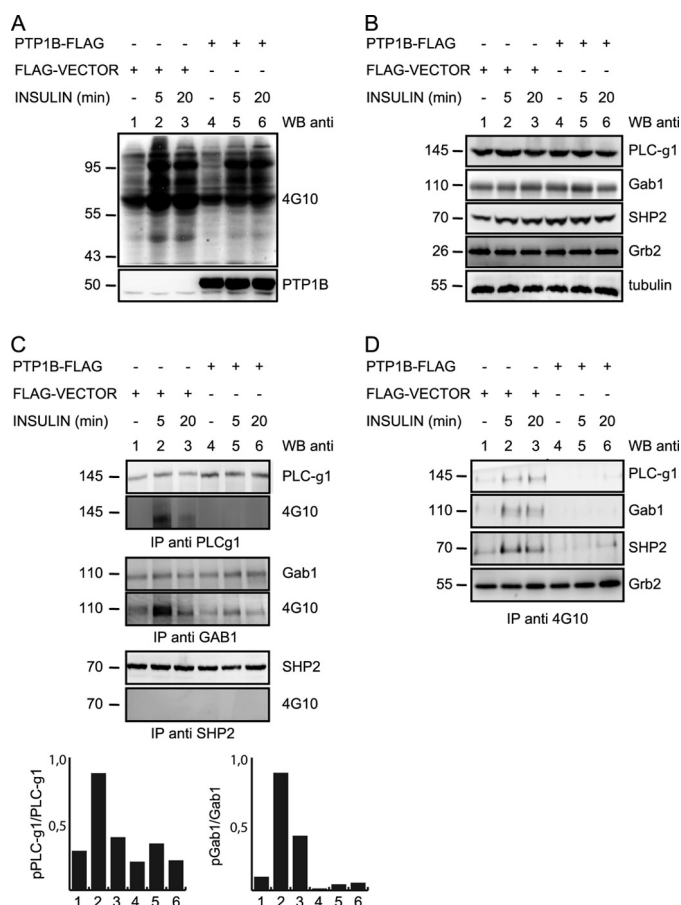


FIGURE 3. PLC- γ 1, Gab1, and SHP2 are dephosphorylated by PTP1B upon insulin stimulation. HEK293 cells were transiently transfected with FLAG-tagged PTP1B (lanes 4–6) and FLAG vector as negative control (lanes 1–3). After serum starvation for 18 h, cells were stimulated with 100 ng/ml of insulin for 5 (lanes 2 and 5) and 20 min (lanes 3 and 6). As control we used unstimulated cells (lanes 1 and 4). *A*, total extracts were revealed with anti-4G10 antibody to monitor tyrosine phosphorylation after insulin induction (WB: anti-4G10) and with PTP1B specific antibody (WB: anti-PTP1B) to confirm phosphatase overexpression in lanes 4–6. WB, Western blot. *B*, the lysates were also revealed with antibodies against PLC- γ 1, Gab1, SHP2, Grb2, and tubulin to detect protein levels in the cell extract. *C*, cell lysates were immunoprecipitated with anti-PLC- γ 1, Gab1, and SHP2 antibodies. The blots were incubated with protein-specific (upper panels) or with anti-phosphotyrosine antibodies (WB: 4G10) to monitor the phosphorylation levels of the immunoprecipitated proteins (lower panels). After normalization of the immunoprecipitated proteins, phosphorylation levels were quantified and plotted. *D*, cell extracts were immunoprecipitated with anti-4G10 and revealed with anti-PLC- γ 1, anti-Gab1, anti-SHP2, and anti-Grb2 antibodies.

the phosphorylation of the inferred targets was due to a direct effect on the targets themselves, rather than a secondary consequence of PTP1B dephosphorylation and inactivation of the insulin or EGF receptor kinases (supplemental Fig. S2).

Depletion of PTP1B Augments PLC- γ 1, Gab1, and SHP2 Tyrosine Phosphorylation—A reduction in the expression of PTP1B should result in an increment in the phosphorylation levels of its specific substrates. Hence, we generated stable cell lines expressing a short hairpin RNA targeting the phosphatase transcript (supplemental Fig. S4) and we knocked PTP1B expression levels down to 20% (Fig. 4A, lanes 3 and 4). To increase tyrosine phosphorylation, we transfected HEK293 with the constitutively active kinase Src Y527F (Fig. 4A, lanes 2 and 4). A fraction of the cell lysates were used to immunoprecipitate the tyrosine-phosphorylated proteins with the

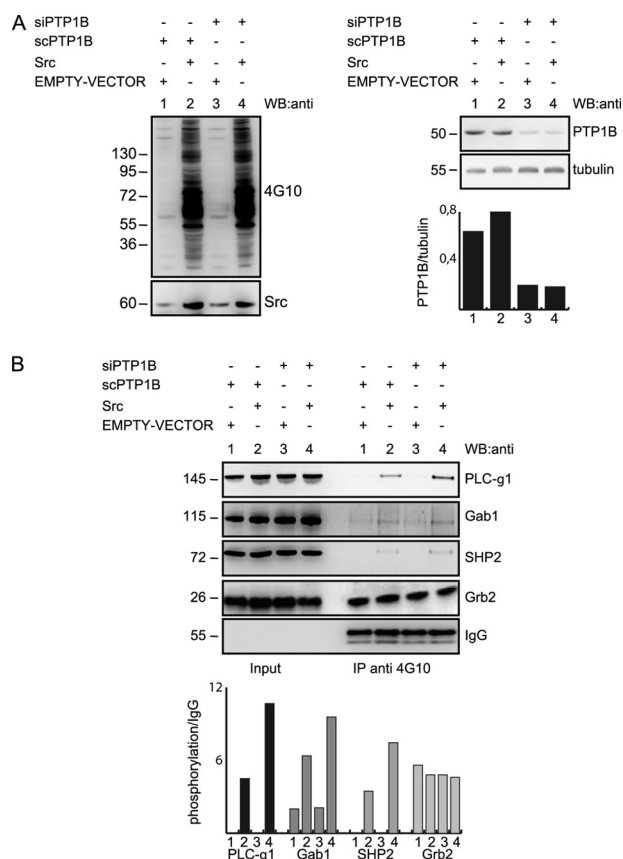


FIGURE 4. Knock down of PTP1B expression by siRNA results in an increase in the level of tyrosine phosphorylation of PLC- γ 1, Gab1, and SHP2. HEK293 stably expressing PTP1B siRNA (lanes 3 and 4) or a scrambled control (lanes 1 and 2) were transiently transfected with a plasmid encoding the constitutively active Src Y527F (lanes 2 and 4) or with the empty vector (lanes 1 and 3). *A*, protein extracts were immunoblotted with the 4G10 antibody, to monitor the phosphotyrosine profile of HEK293 cells, with an antibody against the Src kinase to verify ectopic expression of the constitutively active mutant (lanes 2 and 4), with an anti-PTP1B antibody, to monitor the endogenous PTP1B expression levels and antitubulin, as a loading control. *B*, 1 mg of protein extract was immunoprecipitated with the anti-phosphotyrosine antibody 4G10 (IP: anti-4G10). The input cell extracts and the IP material were immunoblotted (WB) with antibodies against PLC- γ 1, Gab1, SHP2, and Grb2. The amount of IgG was used as loading control. The phosphorylation levels of the indicated proteins were normalized and plotted as a bar diagram.

anti-phosphotyrosine antibody 4G10. The samples were separated by gel electrophoresis and revealed with antibodies against PLC- γ 1, Gab1, SHP2, and Grb2 (Fig. 4B). After normalization of the signal of the immunoprecipitated proteins with the intensity of the immunoglobulin band (IgG), the phosphorylation level in the different conditions was calculated and plotted as a bar diagram. The reduced expression of PTP1B results in a significant increase in PLC- γ 1, SHP2, and Gab1 phosphorylation, when compared with the control cell line expressing the scrambled siRNA (Fig. 4B, compare lanes 2 and 4), whereas consistently with the results in Fig. 3 and supplemental Fig. S2, Grb2 phosphorylation is not affected by PTP1B down-regulation. A similar increase in tyrosine phosphorylation of the inferred targets was observed also after knocking down the expression of PTP1B with independent siRNAs (supplemental Fig. S3). These results indicate that a decrease of PTP1B expression is accompanied by an increase in the phosphorylation of PLC-

γ 1, Gab1, and SHP2 when the phosphorylation is triggered by an activated Src kinase.

PTP1B Directly Dephosphorylates Gab1 and PLC- γ 1—To prove the direct dephosphorylation of the putative PTP1B substrates, Gab1, PLC- γ 1, and Grb2, as a negative control, were immunoprecipitated from cells ectopically expressing a constitutive active Src kinase, to induce tyrosine phosphorylation. When the putative substrates are incubated with purified PTP1B (Fig. 5), the phosphorylation levels of Gab1 and PLC- γ 1 decrease, at increasing incubation times. By contrast, the Grb2 phosphorylation level remains stable after 20 min of incubation with the PTP1B phosphatase.

PTP1B Specificity: the Related Phosphatase TC-PTP Does Not Dephosphorylate PLC- γ 1, Gab1, and SHP2 Proteins—To evaluate PTP1B substrate specificity we used inducible Fln-T-REx-293 cell lines expressing either PTP1B or TC-PTP under the control of doxycycline. PTP1B and TC-PTP are closely related phosphatases displaying a similar domain organization and sharing 70% amino acid sequence identity within their catalytic domains. However, they differ functionally (23) and their substrate specificity, described so far, shows only a limited overlap. The relative expression of PTP1B and TC-PTP was monitored, by probing with an anti-FLAG antibody. TC-PTP is expressed twice as much as PTP1B in these conditions (Fig. 6A). T-REx-293 cells, inducible for either PTP1B (Fig. 6B) or TC-PTP (Fig. 6C), were treated with doxycycline for 24 h, to induce phosphatase expression and stimulated for 5 min with EGF. Affinity purified phosphorylated proteins were immunoblotted with antibodies against PLC- γ 1, Gab1, SHP2, and Grb2 (Fig. 6, B and C). Because the phosphorylation level of Grb2 does not change if either of the two phosphatases are overexpressed, we used the intensity of its signal for normalization of the phosphorylation of the putative substrate proteins (Fig. 6, B and C). Consistent with the results of the experiments so far, as illustrated in the bar diagrams in Fig. 6, we could confirm that PTP1B overexpression causes a decrease of the EGF-induced phosphorylation of PLC- γ 1, Gab1, and SHP2. Conversely, overexpression of the related phosphatase TC-PTP, whereas appreciably reducing the phosphorylation of its validated target Shc (24), has hardly any effect on the inferred PTP1B substrates. These results further confirm the specific role played by PTP1B in dephosphorylation of the predicted substrates.

PTP1B Specifically Interferes with SHP2 Recruitment by Gab1 and with ERK1/2 Activation—It has been suggested that Gab1 mediates ERK1/2 activation through phosphotyrosine-dependent binding and consequent activation of SHP2 in EGF-stimulated cells (25, 26). We have demonstrated here that PTP1B dephosphorylates Gab1. Thus, to lend functional support to our findings, we entertained the hypothesis that the observed modulation of Gab1 phosphorylation levels by PTP1B overexpression could affect the binding of SHP2 to Gab1 and consequently attenuate ERK signaling. To support this hypothesis we transfected HEK293 cells with PTP1B and stimulated them with EGF for 5 min. A fraction of the protein extract was immunoprecipitated with an antibody against Gab1 and, after electrophoresis, it was revealed with antibodies

against Gab1 and SHP2, to monitor formation of a Gab1-SHP2 complex (Fig. 7A). Gab1 phosphorylation after EGF induction was tracked with an anti-phosphotyrosine antibody (4G10) and by far Western with the two SH2 domains of SHP2 (Fig. 7A, left panel) (27). EGF stimulation (lane 3) results in tyrosine phosphorylation of Gab1 and, as a consequence, in formation of a Gab1-SHP2 complex (28). PTP1B overexpression (lane 4) causes a substantial decrease of the amount of SHP2 associated to Gab1, a decrease that parallels a reduction in Gab1 phosphorylation and in SHP2-SH2 binding (Fig. 7A, compare lanes 3 and 4). Concurrently, we observed that ERK activation following EGF induction is markedly reduced when PTP1B is overexpressed (Fig. 7A, central panel, compare lanes 3 and 4). To further investigate the interaction between PTP1B and Gab1, we modulated the levels of PTP1B in a doxycycline inducible 293 cell line where, after induction, the concentration of PTP1B is only three times as high as the endogenous. After 24 h of PTP1B expression, we confirm the dephosphorylation of Gab1 and the decrease of the Gab1-SHP2 complex (Fig. 7B). The observed impact of PTP1B on formation of the Gab1-SHP2 complex is specific because it is not detected in inducible cell lines expressing the related phosphatase TC-PTP (supplemental Fig. S5). Consistently, incubation with a specific dibromohydroxybenzoyl PTP1B inhibitor (29) results in an increase of Gab1 phosphorylation and consequently in an increment of the Gab1-SHP2 complex and MAPK activation (Fig. 7D, compare lanes 3 and 4). At the same time, we observe the formation of an EGF-dependent complex between Gab1 and PTP1B (Fig. 7, C and D, lanes 3 and 4). This latter complex is probably mediated by interaction with the adaptor protein Grb2, which is stably associated with PTP1B via its SH3 domains (Fig. 7C). Several authors reported the formation of a ternary complex between PTP1B and substrates mediated by adaptor proteins such as Grb2 (30) and p130^{cas} (31). This ternary complex probably promotes the dephosphorylation of PTP1B substrates by bridging the enzyme to its targets. Taken together these results establish a strong correlation between the concentration and the activity of PTP1B, its ability to bind (indirectly) and de-phosphorylate Gab1, the dissociation of a complex between Gab1 and SHP2, and the concomitant inactivation of ERK1/2.

DISCUSSION

Post-translational modifications play a crucial role in most biological processes. Regulated addition and removal of chemical groups such as phosphates, methyl, or acetyl groups etc. may change protein activity and modulate their integration in the protein interaction web. The enzymes that modify proteins (e.g. kinases) often show an intrinsic substrate preference. However, the enzymatic selectivity that can be measured *in vitro* cannot account for the substrate specificity observed *in vivo*. Although conserved sequence motifs can be recognized in kinase targets (32), these motifs are not sufficiently informative to confidently identify *in vivo* substrates. Other factors, like protein co-expression, co-localization, and substrate capture by the formation of complexes via non-catalytic domains, may play an important or even a dominant role.

PTP1B Substrates in the Ras Pathway

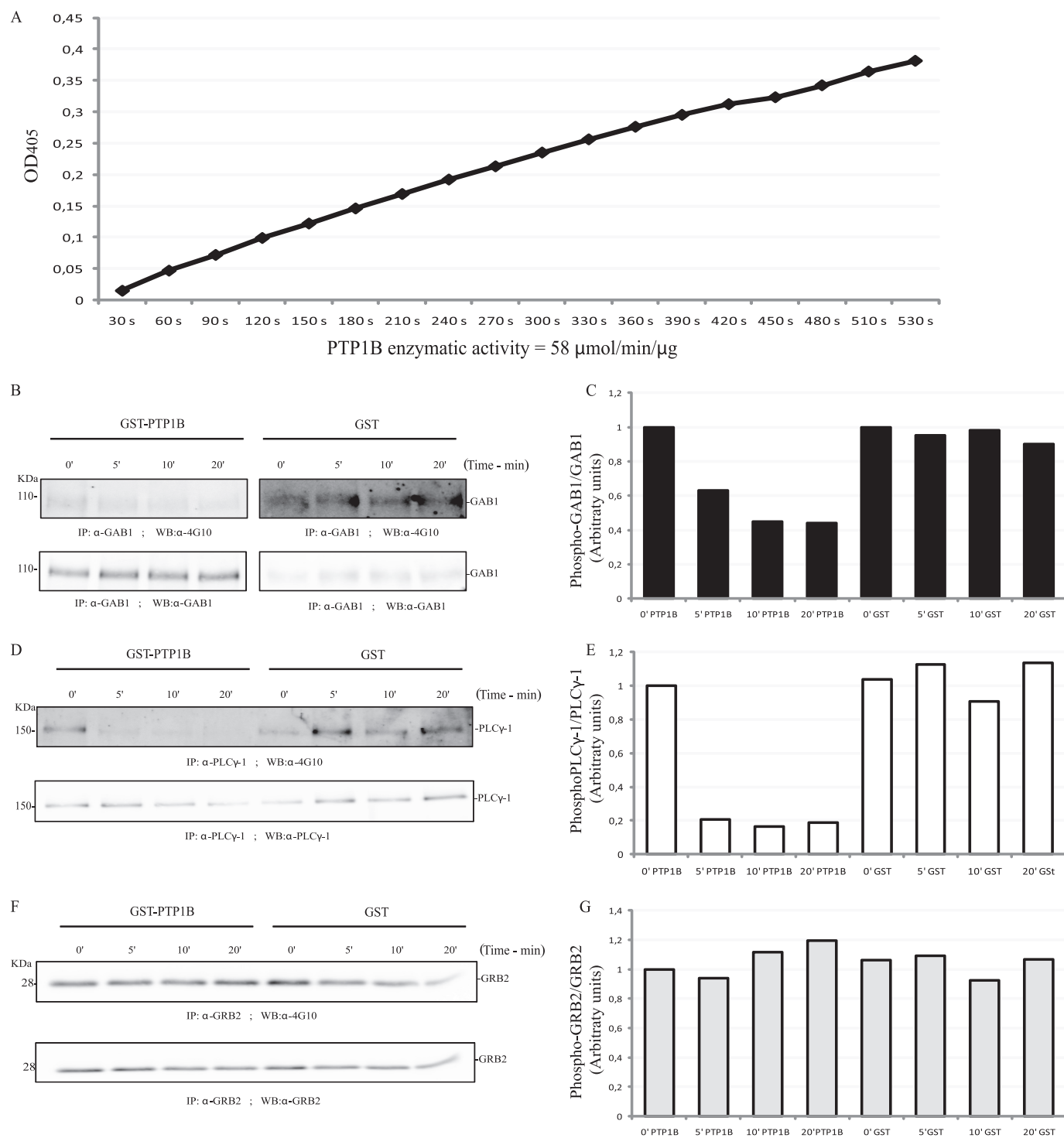


FIGURE 5. PTP1B directly dephosphorylates Gab1 and PLC- γ 1 but not GRB2 *in vitro*. *A*, the PTP1B catalytic domain was expressed in bacteria, purified, and incubated with *p*-nitrophenyl phosphate substrate in assay buffer to demonstrate phosphatase activity. The increase in the absorbance at 405 was plotted as a function of time, and the specific activity was calculated by applying the Lambert-Beer law. *B*, *D*, and *F*, T-Rex-293 cells were transiently transfected with 10 μg of a plasmid directing the synthesis of the Src mutant Y527F, a constitutively active mutant of SRC kinase. After cell lysis the protein extracts were immunoprecipitated with anti-Gab1, anti-PLC γ -1, or anti-Grb2 antibodies. The immunoprecipitated proteins were incubated with 0.3 μg of GST-PTP1B catalytic domain or GST alone, for the indicated times, at 30 $^{\circ}\text{C}$, in reaction buffer. The tyrosine phosphorylation levels and the immunoprecipitated proteins were revealed, respectively, with anti-4G10 and anti-Gab1, PLC γ -1, and Grb2 antibodies (*C*, *E*, and *G*). The band intensities were quantitated and the ratios of phosphoproteins (*p*-Gab1, *p*-PLC γ -1, and *p*-Grb2) versus total proteins (Gab1, PLC γ -1, and Grb2) are represented as bar graphs.

Exploiting this contextual information in predictive strategies can dramatically improve the ability to infer functional targets (7). These considerations are even more important when applied to enzymes such as phosphatases whose sequence selec-

tivity is less pronounced than the one observed in kinases. In this work we have challenged our ability to predict protein-tyrosine phosphatase substrates by a combined experimental and informatic approach.

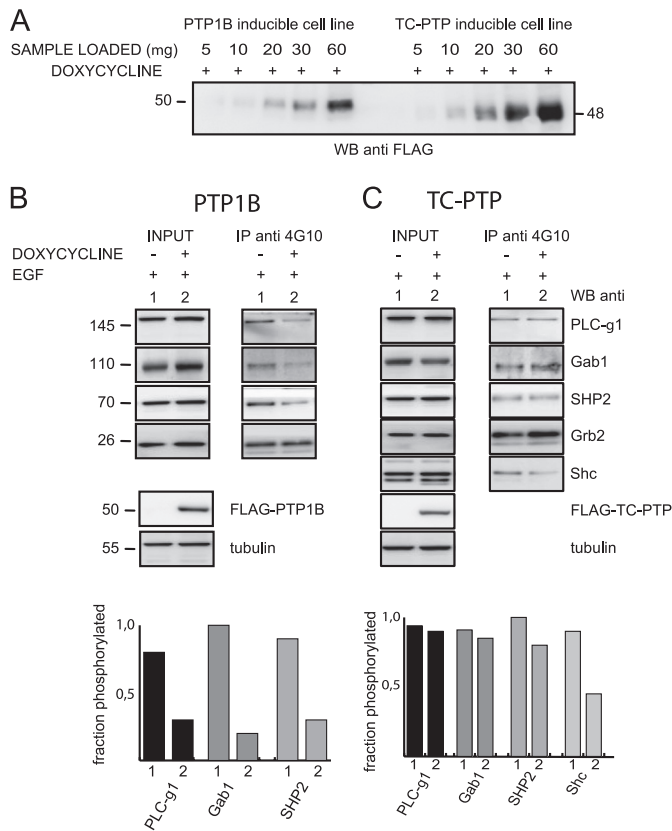


FIGURE 6. The TC-PTP phosphatase does not dephosphorylate PTP1B substrates. T-REx-293 cells were incubated for 24 h with doxycycline to induce the expression of PTP1B (A and B) and TC-PTP (A and C) phosphatases. A, different amounts of total extract expressing both PTP1B and TC-PTP were loaded onto the gel and the blot was revealed with anti-FLAG antibody. B and C, cells were stimulated with EGF for 5 min before cell lysis. 1 mg of the protein extract of each sample was immunoprecipitated with 4G10 antibody. The input material (*Input*) and the phosphorylated proteins (*IP anti-4G10*) were immunoblotted with antibodies against the indicated proteins. The phosphorylation levels of the proteins were normalized with the phospho-Grb2 signal and plotted as a bar diagram.

PTPs participate in maintaining the steady state balance of protein tyrosyl phosphorylation, thereby regulating most biological processes. Thus a complete characterization of physiological phosphatase substrates is the key to understanding the molecular mechanisms modulated by this enzyme class. Over the past decades several focused studies have contributed to shed light on the basis of substrate selection. Until recently, no general proteome wide approach aimed at the description of most of the substrates of a given phosphatase was proposed. While our project was in progress, Mertins and co-workers (33) reported a new strategy taking advantage of high resolution LC-MS and differential labeling to identify tyrosines that are hyperphosphorylated in fibroblasts knocked out in the *PTP1B* gene. This approach led to the identification of 17 PTP1B targets. However, perturbation of the phosphotyrosine network by removal of an enzyme as central as PTP1B could also indirectly enhance or down-regulate phosphotyrosine-mediated signaling downstream of direct substrates. In addition, false negatives could arise because of low signal in the MS spectrum or compensatory mechanisms upon PTP1B knock-out, such as up-regulation of functionally related phosphatases complementing the PTP1B defects.

Here we have described a different and complementary approach that combines, in a statistically meaningful model, two features impinging on substrate choice.

PTP1B Substrate Specificity—Although PTPs are considered rather promiscuous enzymes with little preference for specific residues flanking the phosphorylated tyrosine, several approaches have contributed to identify motifs in aligned substrate sequences. A practical strategy to investigate the *in vitro* substrate specificity of PTP enzymes makes use of a mutant form of this enzyme class, dubbed “substrate trapping mutant” (12). These mutants represent a practical tool because they allow conversion of a relatively laborious enzymatic assay into a convenient binding assay that has a high-throughput potential. Here we have combined the versatility of trapping mutants with a new approach that permits to probe protein binding on high density peptide arrays to determine the recognition specificity of PTP1B, one of the best studied members of this enzyme class. The results of this approach confirm, at a much higher resolution, and extend previous observations attesting that the new technology is not introducing major artifacts (18, 34, 35). In addition, the large number of peptides tested has permitted to compute meaningful residue frequencies in peptides that either do or do not bind. We have used these positive and negative datasets to build a PSSM that combines the positional amino acid frequencies in both the positive and negative sets and we have used it to rank any candidate tyrosine-containing peptide according to the probability of being a phosphatase substrate. The resulting PSSM benefits from the unique characteristics of our experimental approach that returns both a list of positive and negative peptide substrates. To assess the impact of taking into account also non-interacting peptides, we evaluated the performance of a PSSM that uses a random selection of peptides as negative dataset or does not use any “negative” information at all. In both cases the AUC was much lower (0.81/0.69) than that of the PSSM that takes into account both positive and negative information (0.91). We also assessed the performance of our predictor by measuring the AUC of the ROC curve obtained using the 15 substrates proposed by the approach of Mertins *et al.* (33) as a source of true positives. In this case the AUC was found to be 0.85.

The PSSM derived from the chip approach has a high resolution and permits substrate inference down to single residue detail. In fact the potential of our PSSM is also attested by the substantial correlation between the PSSM score and the enzymatic (*i.e.* dephosphorylation) activity measured *in vitro* by an independent assay (21) (Fig. 2D). It remains to be established to what extent this correlation would extend to *in vivo* activity. However, given the lack of suitable reagents or methodology to experimentally probe in a quantitative manner each phosphotyrosine peptide, in the Bayesian inference, we have not exploited this opportunity and our predictions are at the protein level.

Bayesian Integration of Orthogonal Features—The “*in vitro*” PTPs substrate recognition specificity, as determined by the peptide chip profiling experiment and encoded in the PSSM, only captures the preference of the enzymatic pocket for specific sequence motifs and as such it may turn out to be insuffi-

PTP1B Substrates in the Ras Pathway

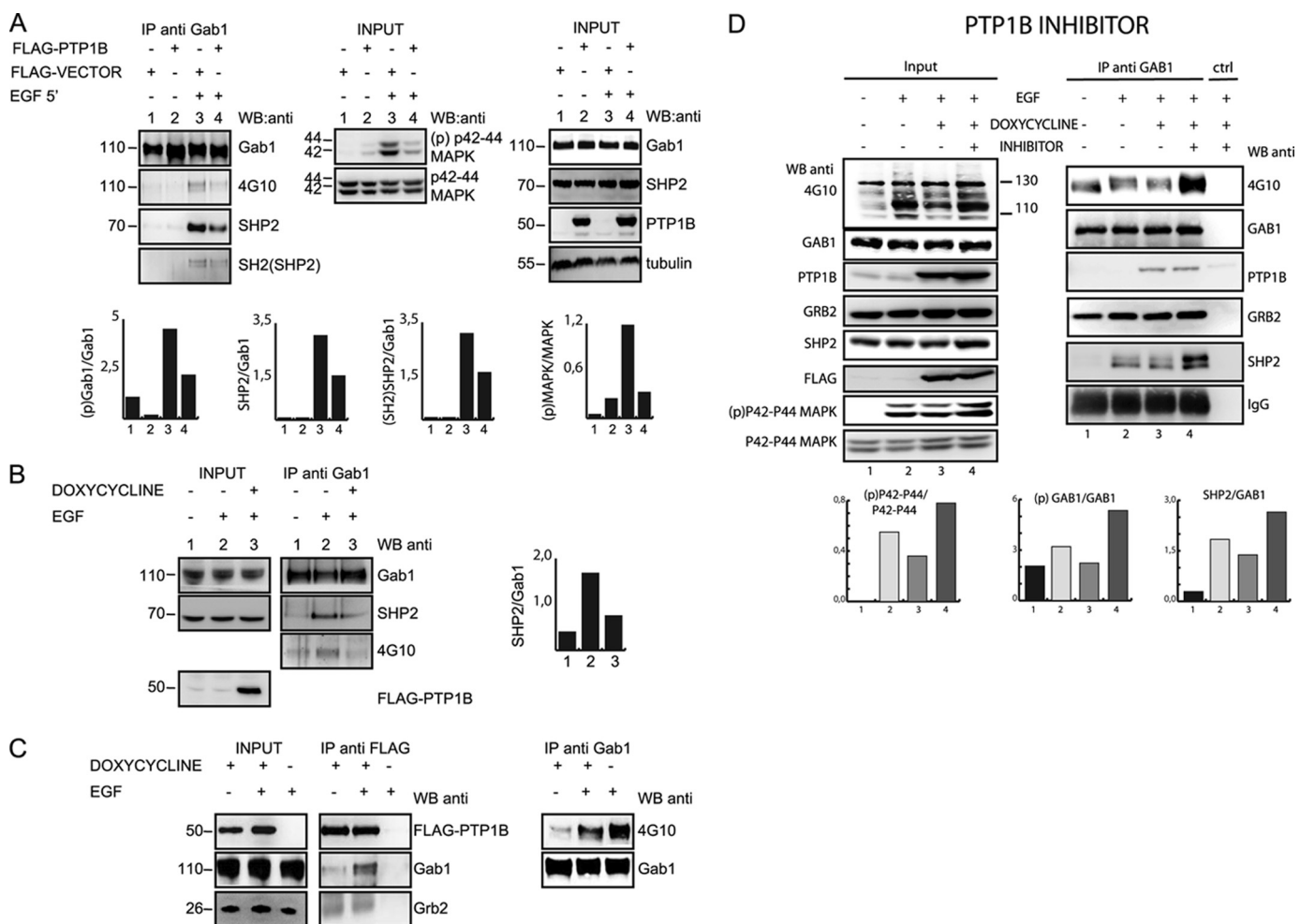


FIGURE 7. PTP1B negatively affects SHP2 recruitment by phosphorylated Gab1 and ERK phosphorylation. *A*, HEK293 cells were transiently transfected with FLAG-tagged PTP1B (lanes 2 and 4) or with the corresponding amount of empty FLAG vector (lanes 1 and 3) and stimulated with EGF for 5 min (lanes 3 and 4). 1 mg of protein extract was immunoprecipitated with anti-Gab1 antibody. The IP material was immunoblotted with antibodies against Gab1 (to control Gab1 immunoprecipitation), 4G10 (to monitor Gab1 phosphorylation level), SHP2 (to verify its interaction with Gab1), and the purified SH2 domains of SHP2 (to verify the interaction with the phosphotyrosines of Gab1). The amount of immunoprecipitated Gab1 was used as a reference to quantify the fraction of Gab1 phosphorylated in the four different experimental conditions. In the bar diagram in *A*, we have quantified the amount of phospho-Gab1 (to verify EGF dependent phosphorylation), the amounts of SHP2 and SH2 domains bound to Gab1 (to quantify the amount of SHP2 bound to phospho-Gab1 in lane 3 and the decrease of interaction due to Gab1 dephosphorylation by PTP1B in lane 4). The *input* was revealed with antibodies against Gab1, SHP2, PTP1B, and tubulin. The anti- β -tubulin antibody was used as a loading control and the anti-PTP1B antibody was used to verify PTP1B transfection. The protein extracts were revealed also with anti-phospho-p42–44 MAPK-ERK1/2, to control the EGF-induced activation of ERK1/2 (lane 3) and the decreased ERK phosphorylation subsequent to PTP1B expression (lane 4). The ratio of phosphorylated p42–44 MAPK to total ERK1/2 protein is plotted. *B*, T-Rex-293 cells were incubated for 24 h with doxycycline to induce the expression of PTP1B phosphatase (lanes 3). Cells were induced 5 min with EGF (lanes 2 and 3). 1 mg of the protein extracts was immunoprecipitated with Gab1 antibody. The *input* and the Gab1-bound proteins (*IP material*) were immunoblotted with antibodies against Gab1, SHP2, and 4G10 (left). The expression of PTP1B was revealed with anti-FLAG antibody (lane 3). Bar diagram in *B*, the amount of immunoprecipitated Gab1 was used as a reference to quantify the SHP2 fraction bound to phosphorylated Gab1 (lane 2) to compare with the SHP2 bound to Gab1 in PTP1B overexpression conditions (lane 3). *C*, to check the co-immunoprecipitation of phospho-Gab1 when SHP2 is immunoprecipitated in the presence of low level of PTP1B, T-Rex-293 cells were incubated for 24 h with doxycycline to induce expression of PTP1B (lanes 1 and 2). Five minutes before cell lysis, cells were stimulated with EGF (lanes 2 and 3). 1 mg of the protein extract of each sample was immunoprecipitated with anti-FLAG antibody (for PTP1B) and 1 mg was used to immunoprecipitate Gab1 protein. The *input material* (*Input*) as well as the IP with anti-FLAG (PTP1B) antibody were immunoblotted with antibodies against FLAG, Gab1, and Grb2. The IP with anti-Gab1 antibody were revealed with 4G10 antibody to detect the phosphorylation level of Gab1. *D*, inhibition of PTP1B activity increases ERK phosphorylation and improves co-immunoprecipitation of Gab1 with SHP2, although not affecting its co-immunoprecipitation with Grb2. T-Rex-293 cells were incubated for 24 h with doxycycline to induce the expression of PTP1B (lanes 3 and 4) and serum starved for 20 h. An aliquot of 293 cells were treated with 250 μ M of the PTP1B inhibitor (Calbiochem compound N0539741) for 1 h (lane 4) and induced with EGF for 5 min (lanes 2–4). Protein extracts were revealed with antibodies against the indicated proteins. The protein extracts were also immunoblotted with anti-phospho-p42–44 MAPK-ERK1/2, to monitor the activation of ERK1/2. 1 mg of protein extract was immunoprecipitated with an anti-Gab1 antibody and revealed with 4G10 antibodies (to monitor Gab1 phosphorylation levels), anti-Gab1 (to control Gab1 immunoprecipitation), and anti-SHP2, to verify its interaction with Gab1. The fractions of phosphorylated p42–44 MAPK and phosphorylated Gab1 are plotted as bar diagrams together with the fraction of SHP2 bound to Gab1.

cient for the discovery of new substrates of potential physiological relevance. Contextual information, such as protein localization and expression, or the coordinated operation of adapter proteins mediating the formation of molecular complexes may come into play to direct the phosphatase toward

its enzymatic targets (32). For this reason, we sought to combine our experimentally derived PSSM with network context information encoded in the WID matrix as the shortest graph distance between any two proteins. Given the orthogonality of the two pieces of evidence, we adopted a naive Bayes ap-

proach to build an integrated model combining the two features. The combined model outperforms both models constructed by singling out one feature at a time, thus establishing that the information contained in the PSSM and WID matrix complement each other and both contribute to increase predictions reliability. A third feature that we considered in our Bayesian model is protein co-expression, as inferred from a gene coexpression table obtained by compiling the results of a large collection of gene chips experiments (36). We eventually did not consider this feature because it performed only slightly better than random in classifying the PTP1B substrates in the golden standard list (AUC = 0.61) and it did not contribute to improve the model based on the PSSM and the WID matrix.

Despite the good performance of our approach, caution is recommended. Both the collection of naturally occurring human tyrosine phosphorylation sites and the human protein-protein interaction network are largely incomplete. Thus, substrates may be missed because no experimental evidence has been produced to accurately integrate them in the protein interaction network or their phosphorylation has not been reported. In addition, the AUC of the ROC evaluating the predictive power of WID might be overestimated. Although we took the precaution of removing from the WID matrix all the information reporting direct interactions between PTP1B and the proteins in our substrate golden standard, we still feel that this feature could be affected by a "social bias." Once an interesting substrate is reported in the literature, the community is motivated to analyze possible interactions with proteins that are in the functional/physical neighborhood. This attitude could influence the density of the (indirect) connections linking PTP1B and its described substrates, thereby decreasing their graph distance in the network as compared with substrates yet to be discovered. Conversely, a poorly characterized phosphatase will not benefit from this feature because of its limited integration in WINT_homo. Notwithstanding these caveats, our results indicate that the approach that we have reported here can lead to the discovery of new PTP1B substrates and may be extended to other phosphatases.

The PTP1B Network—Because of its involvement in modulating the outcome of numerous human disorders, PTP1B is the most intensively studied tyrosine phosphatase. Despite the wealth of available information, the network of interactions, and enzymatic substrates that have been assembled over the years is not sufficient to fully explain the physiological outcome of modulating the levels of this phosphatase in cells or living organisms. A variety of reports are consistent with the notion that the main role of PTP1B is to act as a brake for proliferative and metabolic signals (37, 38). However, more recent data indicate that this simple model is somewhat limited. In fibroblast, PTP1B is required for the activation of the small GTPase Ras, an enzyme that is generally associated with increased cell proliferation and motility (39). Consistent with a positive role in signaling, PTP1B has been identified as the major tyrosine phosphatase that positively regulates the activity of endogenous Src kinase by reducing phosphorylation at Tyr-527 both in colon (40) and in

breast cancer (41, 42). These and other somewhat contradictory reports indicate that we are far from a system level understanding of the PTP1B network. Because the best characterized role of PTP1B is in the regulation of proliferative pathways, our *in vivo* validation experiments focused on proteins that are involved in modulating the response to receptor tyrosine kinase stimulation: PLC- γ 1, SHP2, Gab1, EGFR, and SHP1.

Using the substrate-trapping approach, we demonstrate a specific and direct binding of the inferred substrate proteins to the catalytic domain of PTP1B (Fig. 2). Consistently PTP1B overexpression down-regulates the phosphorylation of PLC- γ 1, Gab1, and SHP2, after insulin or EGF stimulation (Fig. 3). Because it has been reported that PTP1B can act as a negative regulator of these pathways by directly de-phosphorylating the receptors (43), we have also monitored substrate phosphorylation levels under phosphorylation conditions that are independent from growth factor receptor activity. By making use of a constitutively active, and PTP1B-insensitive form of the Src kinase, we lend further support to a direct dephosphorylation mechanism and exclude an indirect effect due to the inactivation of upstream kinases (*i.e.* the receptor). The results obtained by knocking down the PTP1B concentration by an siRNA approach or by abating its enzymatic activity with a specific inhibitor are consistent with these conclusions (Figs. 4 and 7). The observed de-phosphorylation of PLC- γ 1, SHP2, and Gab1 is specific because the phosphorylation level of the adaptor protein Grb2 is found to be independent of PTP1B concentration. At the same time, increasing the concentration of TC-PTP, the closest PTP1B homologue, does not affect the phosphorylation levels of the proposed PTP1B substrates, thus confirming some level of specificity in PTP1B substrate selection (Fig. 6 and supplemental Fig. S5). Direct dephosphorylation of Gab1 and PLC- γ 1 by PTP1B is proved by an *in vitro* protein-tyrosine phosphatase assay with purified enzyme and immunoprecipitated substrates (Fig. 5). PLC- γ 1 and SHP2 were also in a list of putative PTP1B substrates recently reported by Mertins *et al.* (33), whereas Gab1 is proposed and validated as a PTP1B substrate for the first time here, to our knowledge. Gab1 is a member of a family of docking/scaffolding proteins acting downstream of tyrosine kinase receptors. After EGFR stimulation, Gab1 is phosphorylated on several tyrosines that in turn work as docking sites for SH2 domain-containing signaling molecules, such as SHP2, PI3K regulatory subunit p85, Crk, Ras GTPase-activating protein, and PLC- γ 1 (some of these interactions are shown in Fig. 8). We asked whether the PTP1B-dependent modulation of Gab1 tyrosine phosphorylation described here could affect its ability to form signaling complexes. Among the many Gab1 binding partners, we focused on SHP2, a ubiquitously expressed protein-tyrosine phosphatase, which, in association with Gab1, has a crucial role in the receptor tyrosine kinase-dependent activation of ERK1/2 (25). Here we have shown that PTP1B can down-regulate the formation of the SHP2-Gab1 complex by dephosphorylating Gab1 (Fig. 7). We also provide

PTP1B Substrates in the Ras Pathway

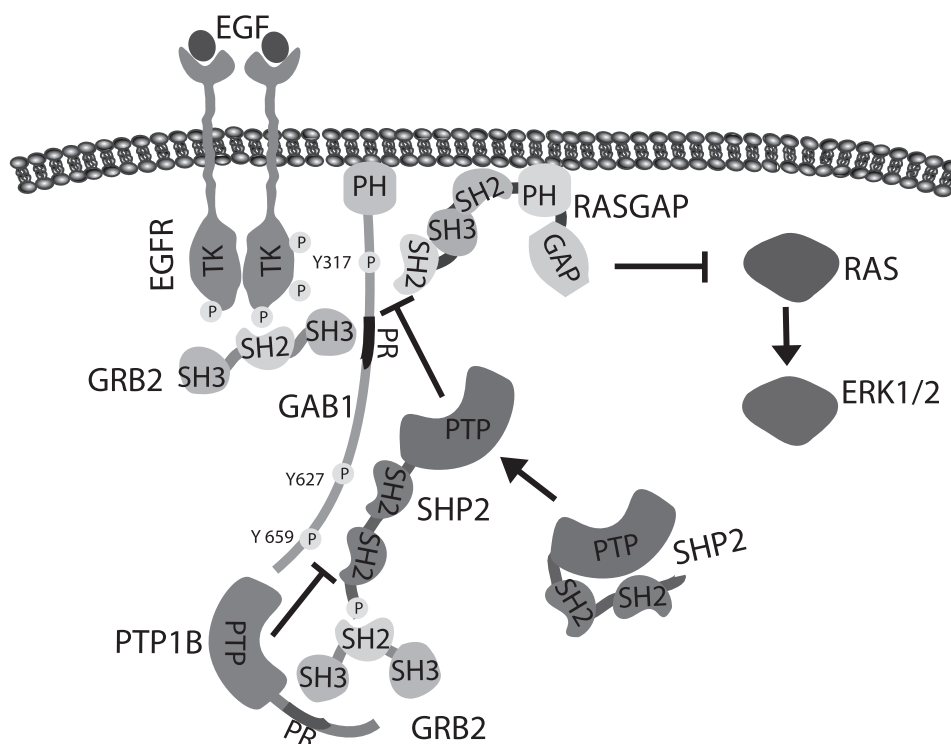


FIGURE 8. Model of the proposed mechanism by which PTP1B controls ERK activity. The schematic only shows the proteins and the interactions that are relevant for the novel mechanism described in this article. In principle, PTP1B could indirectly associate with Gab1 via interactions mediated by the two SH3 domains of Grb2. However, we have chosen to show a complex that links PTP1B to Gab1 via two intermediate interactions mediated by the SH2 domains of Grb2 and SHP2. Although we have no direct proof of the existence of such a tetraplex, such an interaction topology, based on established binary interactions can better explain our observation that the association between PTP1B and Gab1 is EGF dependent. The SHP2 phosphatase is shown to activate ERK by dephosphorylating Gab1 phosphotyrosine 317 thus inhibiting formation of a Gab1-RasGAP complex (45). Here we add that this positive regulation of the ERK pathway can be shut down by the PTP1B-mediated dephosphorylation of the SHP2 docking sites on Gab1 (Tyr-627 and Tyr-659). *PR* stands for proline rich, *PH* for pleckstrin homology domain, *SH* for Src homology domain, *TK* for tyrosine kinase, and *PTP* for phosphotyrosine phosphatase domain.

evidence of a physical association between Gab1, PTP1B, and the docking protein Grb2 to form a ternary complex. PTP1B has been shown to have the potential to interact directly with Grb2 via a proline-rich motif that docks into the Grb2 SH3 domains (31). We confirmed this interaction demonstrating a considerable co-immunoprecipitation of these two proteins. More specifically we showed a constitutive interaction between PTP1B and Grb2 both in non-stimulated and EGF-stimulated cells, whereas the association of PTP1B with Gab1 was observed only after EGF induction. As mentioned before, the interaction between Grb2 and PTP1B and the functional effects on Gab1 dephosphorylation described in the present work have implications for the regulation of proliferation. The role of PTP1B in the Ras/MAP kinase pathway is controversial. Most studies have been performed in immortalized fibroblast cell lines derived from wild type and PTP1B knockout embryos stimulated with PDGF. Although many studies point to the role of PTP1B as a growth suppressor, Haj and co-workers (10) reported that PDGF-stimulated ERK and AKT activation is not substantially altered in PTP1B-deficient mouse embryonic fibroblasts. However, in the same cellular system under lower levels of PDGF stimulation (10–20 ng/ml), the absence of PTP1B was shown to attenuate ERK activation (39). In our work we used a HEK293 cell line under stimulation at high EGF concentration and we observed that overexpression of PTP1B re-

duces ERK activation, whereas inhibition of the PTP1B enzymatic activity has an opposite effect. Taken together, these results point to a subtle control mechanism heavily dependent on the cellular and genetic context. Our results lend support to a new mechanism whereby PTP1B interferes with the association of Gab1 with SHP2 in response to EGF treatment (Fig. 8). The PTP1B-triggered dissociation of SHP2 mimics the Leopard syndrome mutants of p120Ras GAP, mediated by the attachment of its SH2 domain to the phosphorylated Tyr-317 of Gab1. Consistently, we observe that this rearrangement of the Gab1 complex parallels a drastic decrease in ERK activation. Given the intricacy of the network, this does not conclusively prove that dissociation of the SHP2-Gab1 complex is the only main route leading to the PTP1B-mediated ERK inactivation. In conclusion, these data suggest that PTP1B may play an important role in regulating receptor tyrosine kinase signaling pathways, by modulating the ability of Gab1 to form complexes with SH2 domain-containing proteins such as the phosphatase SHP2 and, consequently, with RasGAP and p85 PI3K.

REFERENCES

1. Pawson, T., and Scott, J. D. (2005) *Trends Biochem. Sci.* **30**, 286–290
2. Stuiblé, M., Doody, K. M., and Tremblay, M. L. (2008) *Cancer Metastasis Rev.* **27**, 215–230
3. Andersen, J. N., Jansen, P. G., Echwald, S. M., Mortensen, O. H., Fukada,

- T., Del Vecchio, R., Tonks, N. K., and Møller, N. P. (2004) *FASEB J.* **18**, 8–30
4. Tonks, N. K. (2006) *Nat. Rev. Mol. Cell Biol.* **7**, 833–846
 5. Tiganis, T., and Bennett, A. M. (2007) *Biochem. J.* **402**, 1–15
 6. Easty, D., Gallagher, W., and Bennett, D. C. (2006) *Curr. Cancer Drug Targets* **6**, 519–532
 7. Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A., Jørgensen, C., Miron, I. M., Diella, F., Colwill, K., Taylor, L., Elder, K., Metalnikov, P., Nguyen, V., Pasculescu, A., Jin, J., Park, J. G., Samson, L. D., Woodgett, J. R., Russell, R. B., Bork, P., Yaffe, M. B., and Pawson, T. (2007) *Cell* **129**, 1415–1426
 8. Ostman, A., and Böhmer, F. D. (2001) *Trends Cell Biol.* **11**, 258–266
 9. Dadke, S., Kusari, A., and Kusari, J. (2001) *Mol. Cell Biochem.* **221**, 147–154
 10. Haj, F. G., Markova, B., Klamann, L. D., Bohmer, F. D., and Neel, B. G. (2003) *J. Biol. Chem.* **278**, 739–744
 11. Liu, F., and Chernoff, J. (1997) *Biochem. J.* **327**, 139–145
 12. Blanchetot, C., Chagnon, M., Dubé, N., Hallé, M., and Tremblay, M. L. (2005) *Methods* **35**, 44–53
 13. Diella, F., Cameron, S., Gemünd, C., Linding, R., Via, A., Kuster, B., Sicheritz-Pontén, T., Blom, N., and Gibson, T. J. (2004) *BMC Bioinformatics* **5**, 79
 14. Blom, N., Gammeltoft, S., and Brunak, S. (1999) *J. Mol. Biol.* **294**, 1351–1362
 15. Flint, A. J., Tiganis, T., Barford, D., and Tonks, N. K. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1680–1685
 16. Vacic, V., Iakoucheva, L. M., and Radivojac, P. (2006) *Bioinformatics* **22**, 1536–1537
 17. Garaud, M., and Pei, D. (2007) *J. Am. Chem. Soc.* **129**, 5366–5367
 18. Sarmiento, M., Puius, Y. A., Vetter, S. W., Keng, Y. F., Wu, L., Zhao, Y., Lawrence, D. S., Almo, S. C., and Zhang, Z. Y. (2000) *Biochemistry* **39**, 8171–8179
 19. Milarski, K. L., Zhu, G., Pearl, C. G., McNamara, D. J., Dobrusin, E. M., MacLean, D., Thieme-Seffler, A., Zhang, Z. Y., Sawyer, T., Decker, S. J., et al. (1993) *J. Biol. Chem.* **268**, 23634–23639
 20. Schreiber, T. B., Mäusbacher, N., Breitkopf, S. B., Grundner-Culemann, K., and Daub, H. (2008) *Proteomics* **8**, 4416–4432
 21. Barr, A. J., Ugochukwu, E., Lee, W. H., King, O. N., Filippakopoulos, P., Alfano, I., Savitsky, P., Burgess-Brown, N. A., Müller, S., and Knapp, S. (2009) *Cell* **136**, 352–363
 22. Ceol, A., Chatr Aryamontri, A., Licata, L., Peluso, D., Briganti, L., Perfitto, L., Castagnoli, L., and Cesareni, G. (2010) *Nucleic Acids Res.* **38**, D532–539
 23. Bourdeau, A., Dubé, N., and Tremblay, M. L. (2005) *Curr. Opin. Cell Biol.* **17**, 203–209
 24. Tiganis, T., Bennett, A. M., Ravichandran, K. S., and Tonks, N. K. (1998) *Mol. Cell Biol.* **18**, 1622–1634
 25. Meng, S., Chen, Z., Munoz-Antonia, T., and Wu, J. (2005) *Biochem. J.* **391**, 143–151
 26. Cai, T., Nishida, K., Hirano, T., and Khavari, P. A. (2002) *J. Cell Biol.* **159**, 103–112
 27. Shi, Z. Q., Yu, D. H., Park, M., Marshall, M., and Feng, G. S. (2000) *Mol. Cell Biol.* **20**, 1526–1536
 28. Cunnick, J. M., Mei, L., Doupnik, C. A., and Wu, J. (2001) *J. Biol. Chem.* **276**, 24380–24387
 29. Wiesmann, C., Barr, K. J., Kung, J., Zhu, J., Erlanson, D. A., Shen, W., Fahr, B. J., Zhong, M., Taylor, L., Randal, M., McDowell, R. S., and Hansen, S. K. (2004) *Nat. Struct. Mol. Biol.* **11**, 730–737
 30. Goldstein, B. J., Bittner-Kowalczyk, A., White, M. F., and Harbeck, M. (2000) *J. Biol. Chem.* **275**, 4283–4289
 31. Liu, F., Hill, D. E., and Chernoff, J. (1996) *J. Biol. Chem.* **271**, 31290–31295
 32. Miller, M. L., Jensen, L. J., Diella, F., Jørgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S. A., Bordeaux, J., Sicheritz-Pontén, T., Olhovskiy, M., Pasculescu, A., Alexander, J., Knapp, S., Blom, N., Bork, P., Li, S., Cesarani, G., Pawson, T., Turk, B. E., Yaffe, M. B., Brunak, S., and Linding, R. (2008) *Sci. Signal* **1**, ra2
 33. Mertins, P., Eberl, H. C., Renkawitz, J., Olsen, J. V., Tremblay, M. L., Mann, M., Ullrich, A., and Daub, H. (2008) *Mol. Cell Proteomics* **7**, 1763–1777
 34. Espanel, X., Huguenin-Reggiani, M., and Hooft van Huijsduijnen, R. (2002) *Protein Sci.* **11**, 2326–2334
 35. Vetter, S. W., Keng, Y. F., Lawrence, D. S., and Zhang, Z. Y. (2000) *J. Biol. Chem.* **275**, 2265–2268
 36. Obayashi, T., Hayashi, S., Shibaoka, M., Saeki, M., Ohta, H., and Kinoshita, K. (2008) *Nucleic Acids Res.* **36**, D77–82
 37. LaMontagne, K. R., Jr., Flint, A. J., Franza, B. R., Jr., Pandergast, A. M., and Tonks, N. K. (1998) *Mol. Cell Biol.* **18**, 2965–2975
 38. LaMontagne, K. R., Jr., Hannon, G., and Tonks, N. K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14094–14099
 39. Dubé, N., Cheng, A., and Tremblay, M. L. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1834–1839
 40. Zhu, S., Bjorge, J. D., and Fujita, D. J. (2007) *Cancer Res.* **67**, 10129–10137
 41. Arias-Romero, L. E., Saha, S., Villamar-Cruz, O., Yip, S. C., Ethier, S. P., Zhang, Z. Y., and Chernoff, J. (2009) *Cancer Res.* **69**, 4582–4588
 42. Bjorge, J. D., Pang, A., and Fujita, D. J. (2000) *J. Biol. Chem.* **275**, 41439–41446
 43. Romsicki, Y., Reece, M., Gauthier, J. Y., Asante-Appiah, E., and Kennedy, B. P. (2004) *J. Biol. Chem.* **279**, 12868–12875
 44. Edouard, T., Combier, J. P., Nédélec, A., Bel-Vialar, S., Métrich, M., Conte-Auriol, F., Lyonnet, S., Parfait, B., Tauber, M., Salles, J. P., Lezoualc’h, F., Yart, A., and Raynal, P. (2010) *Mol. Cell Biol.* **30**, 2498–2507
 45. Montagner, A., Yart, A., Dance, M., Perret, B., Salles, J. P., and Raynal, P. (2005) *J. Biol. Chem.* **280**, 5350–5360