

BBAMCR 13206

Minireview

A thousand and two protein tyrosine phosphatases

David A. Pot¹ and Jack E. Dixon

Department of Biological Chemistry, The University of Michigan Medical School, Ann Arbor, MI (USA)

(Received 1 April 1992)

Key words: Protein phosphorylation; Tyrosine phosphorylation; Tyrosine dephosphorylation; Protein tyrosine phosphatase

Contents

I. Introduction	35
II. Molecular biology of the PTPases	36
III. Structure and function studies of the PTPases	38
IV. Biology of the PTPases	40
V. Concluding remarks	42
Acknowledgements	42
References	42

I. Introduction

Protein post-translational modification is a method cells use to regulate their response to external and internal stimuli. A common post-translational modification employed in this regulatory mechanism is protein phosphorylation. Protein phosphorylation states can regulate enzyme activity as well as the cellular location of proteins. Cells adjust the phosphate levels of specific proteins through the action of kinases and phosphatases. Phosphorylation-dephosphorylation mechanisms predominantly alter the state of Ser or Thr phosphorylation in proteins. Similar modifications of Tyr have only recently been studied. Although this review will focus predominantly on tyrosine dephosphorylation, it is necessary to briefly review the history of tyrosine phosphorylation. Before 1980 the only regulatory phosphoamino acids identified in proteins were phosphoserine and phosphothreonine. Ser and Thr protein phosphorylation states regulate many metabolic processes (reviewed in Ref. 1). In 1930 a unique catalytic activity was discovered while examining *v-src*, the

protein necessary for neoplastic transformation of chicken fibroblasts by Rous sarcoma virus. Analysis of this protein showed that it could phosphorylate itself and other proteins on tyrosine [2-4]. Analysis of cellular phosphoamino acid content revealed that phosphotyrosine accounts for only 0.01-0.05% of cellular phosphoamino acids in normal tissue culture cells [2]. However, this content increases to 1-3% of total cellular phosphoamino acids upon transformation by oncogenic retroviruses [5]. The increase in tyrosine phosphorylation upon viral transformation suggested that tyrosine kinases may be linked to the removal of growth restraint from cells. Investigations of oncogenes from other tumorigenic viruses showed that many of these proteins also displayed tyrosine kinase activity (reviewed in Ref. 6).

The mechanism of growth factor receptor activation was under investigation at the same time that the discoveries of viral tyrosine kinase activity were being made. Surprisingly, one of the growth factor receptors (epidermal growth factor) was also shown to have tyrosine kinase activity [7]. This identification of a non-viral tyrosine kinase which controls cell growth suggested that tyrosine phosphorylation was important in the physiology of non-transformed cells.

The primary sequence of protein tyrosine kinases had been identified through the cloning and sequencing of viral and cellular oncogenes (reviewed in Ref. 8). Progress in our understanding of tyrosine kinases was rapid and the focus on phosphorylation events domi-

Correspondence to: Department of Biological Chemistry, The University of Michigan Medical School, M541 6 Medical Science I, Ann Arbor, MI 48109-0606.

¹ Present address: Howard Hughes Medical Institute, University of California, San Francisco, 533 Parnassus Ave., U-426, San Francisco, CA 94143-0724, USA.

nated this period of progress. Few investigators were attempting to understand the importance of tyrosine dephosphorylation events (see review in Ref. 9)¹. Progress similar to that made with the tyrosine kinases was not possible with the protein tyrosine phosphatases (PTPase, EC 3.1.3.48) since the pure preparation of PTPase and the corresponding amino acid sequence was not available until 1988 [11] when PTPase 1B was purified from human placenta [12]. The partial amino acid sequence of PTPase 1B [12] yielded two surprises. First, tyrosine-specific phosphatase had no sequence similarity to any of the previously characterized serine and threonine phosphatases, suggesting that PTP 1B was a member of a different family of proteins. Second, a search of known protein sequences revealed that PTPase 1B contained a region of similarity to a major lymphocyte c-ll surface glycoprotein termed CD45. CD45 is a 'receptor-like molecule' present on the lymphocyte cell surface and is composed of a highly glycosylated, differentially spliced extracellular domain, a single transmembrane domain, and a cytosolic domain containing two repeats of a 200-amino-acid sequence showing sequence identity to PTP 1B. At the time this comparison was made, no function for CD45 had been determined, but studies using antibodies directed against its extracellular domain suggested that CD45 was involved in early lymphocyte activation (reviewed in Ref. 13). The similarity between CD45 and PTPase 1B not only suggested a function for the cytoplasmic domain of CD45, but also supported the idea that PTPases possessed a signalling function.

II. Molecular biology of the PTPases

The initial identification of two members of the PTPase family (i.e., PTP 1B and CD45) led to the cloning of numerous other PTPases. These were obtained either through low-stringency screening of cDNA libraries [14–28] or by the use of the polymerase chain reaction (PCR) using oligonucleotide primers designed from conserved amino acid sequences within the PTPases [29–42]. Additional PTPases were identified through sequence similarities to the conserved 'active-site' regions of the PTPases (the 'HCSAGVGR (S/T)G' motif) [20,43–50].

Sequence analysis of these clones showed that they either resembled PTPase 1B or CD45. PTPase 1B is a representative of the 'nonreceptor-like' class of PTPases that have only one copy of the 200-amino-acid conserved PTPase domain and no transmembrane spanning region. Fig. 1 is a schematic diagram of the nonreceptor-like PTPases. The diverse structures of

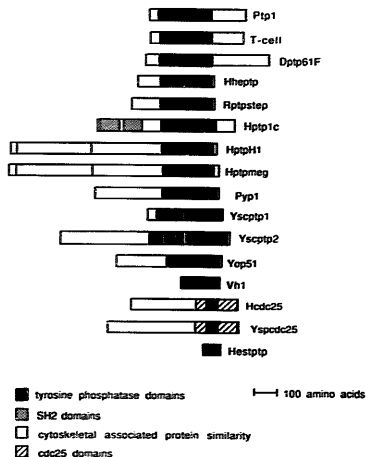


Fig. 1. Primary structures of the nonreceptor-like PTPase family. Names indicated are those first published for the sequence. References: Ptp1 [16,17,29], T-cell [14,26,39], Dptp61F (S. McLaughlin, personal communication), Hheptp [42], Rptpstep [24], Hptp1c [25,40,41], HptpH1 [34], Hptpmeg [33], Pyp1 [31], Yscptp1 [32], Yscptp2 [96], Yop51 [20], Vh1 [49], Hcdc25 [48], Yspcdc25 [44], Hestp1 [50].

these molecules suggest multiple methods of regulation. CD45 is a representative of the 'receptor-like' class of PTPases that have, with two exceptions (PTPase β [23]; and DPTP10D [36,37]), tandem 200-amino-acid PTPase cytosolic repeats and variably sized extracellular domains separated by a single transmembrane spanning region. A review on the structure and potential function of these receptor-like molecules has appeared [51].

The primary structure of 42 PTPases has allowed us to construct a family tree by sequence comparison. This tree, created by performing multiple progressive sequence alignments of each 200-amino-acid conserved region (blackened area in Fig. 1) and clustering related sequences [52], provides a non-biased grouping of the PTPases. The result, shown in Fig. 2, illustrates the relationship of all of the PTPases. In general, the PTPases can be divided into three subfamilies. The receptor-like PTPases form one subfamily while the non-receptor-like class of PTPases can be further divided into additional subfamilies. Approx. 18 nonreceptor-like PTPases group together. PTP1 and PTP 1B are prototypes for this subfamily. A second branch of

¹ For a detailed account of the enzymology of tyrosine phosphatases prior to 1988 the reader is directed to an excellent review by Brautigan [10].

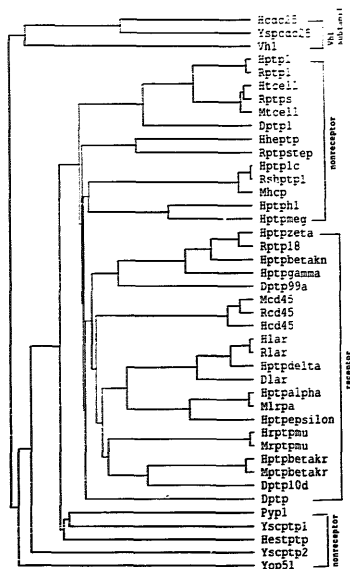


Fig. 2. Family tree of the PTPases. PTPase domains were aligned by progressive alignment method of Feng and Doolittle [52] using the PILEUP program of the Wisconsin Genetics Program version 7.0 [112]. Only PTPase domains and not flanking sequences were input for the analysis. Names of the PTPases are those first published for the sequence. References: Hcdc25 [48], Yspcdc25 [44], Vh1 [49], Dptp1 [17,29], Rptp1 [16], Htcell [14,39], Rptps [26], Mtcell [39], Hptp1 (S. McLaughlin, personal communication), Hheptp [42], Rptpstp [24], Hptp1c [25,40], Rshptp1 [41], Mhpc [40], Hptph1 [34], Hptpmez [33], Hptpzeta [23], Rptp18 [20], Hptpbetakn [21], Hptpgamma [21,23], Dptp99a [36-38], Mcd45 [45], Rcd45 [43], Hcd45 [46], Hlar [47], Rlar [28], Hptpdelta [23], Dlar [15], Hptpalpha [21-23,30], Mirpa [18,19], Hptpepsilon [23], Hptpmu [35], Hptpmu [35], Hptpbetakr [23], Mptpbetakr [35], Dptp10d [36,37], Dptp [15], Pyp1 [31], Yscptp1 [32], Hestptp [50], Yscptp2 [96], Yop51 [20].

this subfamily includes the *vaccinia* virus phosphatase (Vh1) as well as cdc25. Since the *vaccinia* gene encoding the phosphatase was the first member of this subfamily to be identified, we have referred to this grouping as the Vh1 subfamily of phosphatases. It appears that very early in evolution the two nonreceptor-like groups diverged from each other. Interestingly, Vh1 is a phosphatase which catalyzes the removal of phosphate from both serine, threonine, and tyrosine ([49];

Guan et al., unpublished). Results similar to those presented here were described by Krueger et al. using a much smaller database [23].

One difficulty that has arisen due to the rapid cloning of PTPases is the lack of a defined system of nomenclature. For example, there are two different reports of clones named PTPase α , β and γ [21,23]. By chance, both α and γ are identical but, as seen in Fig. 2, a subscript is needed to differentiate the two β molecules. Some discussion of a nomenclature system has taken place at the Annual Protein Phosphatase meeting in Brussels (Sept. 1991). Even though no final system of nomenclature was adopted, a naming system using 'R' as a prefix for receptor-like molecules and no 'R' for nonreceptor-like molecules was suggested. It was further proposed that each PTPase would be given a unique number or letter (or both) assigned in the order cloned. Therefore, with this system, CD45 would be renamed RPTPase 1, and PTPase 1B would be called PTPase 1. No matter what system is chosen, some method of systematic nomenclature is badly needed.

As Fig. 1 indicates, most PTPases contain functionally important sequences other than the conserved 200-amino-acid PTPase domain. By comparison of these sequences with other known protein sequences potential functions can be suggested for these regions of the PTPases. When human LAR was cloned [47], both the intracellular and extracellular domains had no known function. The cytoplasmic domain's similarity to PTPase 1B suggested a function for this part of the molecule. Streuli and collaborators, upon analysis of the extracellular region of this molecule, noted that it possessed similarities to the ligand-binding region of the neural cellular adhesion molecule, N-CAM. Extracellular regions of N-CAM interact in a homophilic manner to influence neuronal development [53]. Since LAR is expressed on the surface of all solid type tissues and contains within its cytoplasmic domain a phosphatase region, this molecule is proposed to be responsible for the contact inhibition of cell growth [15]. Homophilic interactions between LAR molecules located on adjacent cells may cause an increase in LAR's phosphatase activity, a decrease cellular phosphotyrosine content and may thus stop cell growth. Although this is a particularly attractive suggestion, it should be pointed out that we have a very poor understanding of the function of the extracellular domain of the PTPases.

Another sequence similarity that was discovered within the PTPases is in the SH2 domains of PTP1c [25,40,41]. Research by many groups has shown that these 100-amino-acid regions specifically bind tyrosine-phosphorylated proteins (reviewed in Ref. 54). Analysis by Shen and coworkers has shown that the SH2 domains of PTP1c are functional [25]. This suggests that

PTP1c may use these regions to bind substrates or localize to specific subcellular locations, as suggested for other SH2-containing proteins [55]. Two other PTPases, Hptpmeg and HptpH1, contain a region that shows 20–35% similarity to a region contained in three cytoskeletal proteins, band 4.1, ezrin and talin [33,34]. These regions in these cytoskeletal proteins are necessary for association with other cytoskeletal proteins. Therefore this similarity suggests that the two PTPases, termed Hptpmeg and HptpH1, use these sequences to localize themselves to the cytoskeleton and thereby define specific substrates for these PTPases.

Specific amino acid sequences located near the C-terminus of several of the nonreceptor-like PTPases appear to be important in subcellular localization of the phosphatases. Since these sequences appear to play important roles in the intracellular localization of the PTPases, they likely function to define or limit the substrate specificity of the PTPases. Details of these C-terminal sequences and their function will be described later in this review (see Section IV, Biology of the PTPases).

The gene expression pattern of PTPases also suggests specific function. Many of the PTPases are broadly distributed in various cell types; others, however, have quite specific expression patterns. CD45 has long been known to be expressed only in lymphocytic cells [13]. The striatum enriched phosphatase, as its name implies, is expressed in a highly selective manner in the brain's striatum [24]. Selective expression of three *Drosophila* receptor-like PTPases in embryonal central nervous system axons suggests their role in axon outgrowth and guidance [36–38]. PTPase 1 is strongly expressed in the hippocampus [16], a region of the brain known to be important in long-term potentiation and memory. In-situ staining for PTPase activity in the brain using *o*-phospho-L-tyrosine as a substrate [56] confirmed that the hippocampus is a region of the brain rich in PTPases.

III. Structure and function studies of the PTPases

Expression of cloned PTPases in a variety of prokaryotic and eukaryotic systems has provided sufficient quantities of protein for structure/function studies. Using the large database of PTPase sequences, highly conserved residues within the PTPase family can be identified and mutated to examine their importance in catalysis and stability. Deletion mutagenesis can be used to determine the importance of various regions of the molecule. Finally, kinetic analysis and chemical modification of the PTPases can be used to evaluate their mechanism(s) of catalysis.

Two sequence motifs present in kinases and dehydrogenases appear to be conserved in many of the phosphatases. The first sequence is the GXGXXG

motif in the 'active-site region' [HCSAGVGR(S/T)G] of the PTPases. As noted by Pot et al. and Guan et al. [28,32], this motif is also present in dehydrogenases [57] and protein kinases [58] and is located at or near the active site of these enzymes. Modification of the first glycine in the sequence to any other residue eliminates human LAR PTPase activity [59] suggesting that the motif is critical to phosphate structure or function. The presence of this motif suggests that structural features within this region of the PTPases are similar to structures of either the dehydrogenases or the kinases. It has recently been noted that the GXGXXG motifs are functionally different in dehydrogenases and kinases [60]. It will be interesting to determine the functional role of this motif in the PTPases.

The second conserved motif, that appears in the T-cell and PTPase 1 subfamily of the non-Vh1-like PTPases, is the FKVRES sequence [28]. This sequence is located just upstream from the conserved 'HCSAGVGR(S/T)G' motif of the PTPases. The FLVRES sequence corresponds to the most highly conserved region of the phosphotyrosine-binding SH2 domain of the *Src* tyrosine kinase family [54]. The FLVRES sequence in *c-src* may be associated with cell-specific oncogenesis, as mutation of these residues modifies the transformation pattern of two cell lines from different species [61,62]. It is tempting to speculate that PTPases that contain this sequence, which is invariably located close to the active site of the PTPase, also utilize these amino acids in phosphotyrosine binding.

Because phosphotyrosine content of cells is maximally 3% of total cellular phosphoamino acids, even in transformed cells [5], the natural substrates for the PTPases are rare. Therefore most studies of PTPases use proteins that are readily available in pure form and easily phosphorylated on tyrosine. These include proteins such as casein, histone, myelin basic protein, and lysozyme as well as small peptides such as poly (Glu₄, Tyr), and angiotensin. All substrates are generally phosphorylated on tyrosine in vitro using purified tyrosine kinases. The use of these 'artificial' substrates has made it possible to rapidly analyze the kinetic properties of the PTPases. However, the use of artificial substrates limits our understanding of the 'naturally' occurring substrates of the PTPases. Recently, several studies of the PTPases on naturally occurring substrates have been reported. All but the Vh1 family of phosphatases catalyze the removal of phosphate from tyrosine and not serine- or threonine-containing substrates. Demonstrating serine and threonine phosphatase activity in the Vh1 family of PTPases first documented that this family of enzymes did not have a tyrosine-restricted substrate specificity. In addition, the fact that Vh1 showed sequence identity to cdc25 provided the first suggestion that cdc25 was a phosphatase

which could dephosphorylate p34^{cdc2}. The sequence similarity between the two molecules also provided a rationale for testing the phosphatase activity of cdc25 *in vitro* and *in vivo* [63,64].

The insulin receptor tyrosine kinase has been an attractive 'naturally' occurring substrate for the PTPases. Autophosphorylation of the insulin receptor β subunit specifically in the region surrounding amino acid 1150, activates the insulin receptor to phosphorylate other cellular proteins on tyrosine [65]. The regulatory 1150 region contains three tyrosines, all of which must be phosphorylated before trans-kinase activity is possible. Levine and coworkers have established that soluble insulin receptor kinase phosphorylates a peptide containing these tyrosines in the order 1150, 1151, 1146 [66]. Differences in the dephosphorylation of these tyrosines by four tyrosine phosphatases, human T-cell PTPase, human PTPase 1B, human CD45 and rat LAR, on the same triphosphorylated peptide, have recently been described [67]. The two receptor-like phosphatases, CD45 and LAR, remove the phosphates in the rate-dependent order 1146 > 1151 > 1150, essentially reversing the order of phosphorylation. PTPase 1B removes the phosphates in manner, 1150 = 1151 > 1146. T-cell PTPase displays no order of dephosphorylation, removing all phosphates at the same rate. These data implicate substrate specificity as being an important determinant in PTPase catalysis and regulation of substrate specificity. Kinetic studies by Cho et al. [68] support the observations of Ramachandran and workers [67]. They observed that the first domain of human LAR dephosphorylates phosphotyrosine 1146 at least 25 times more efficiently than tyrosine 1150 and 1151 [68].

Many compounds modulate the activity of protein tyrosine phosphatases. Probably the most universal is orthovanadate, which effectively inhibits all tyrosine phosphatases so far tested at levels between 10 and 100 μ M. Another universal inhibitor is molybdate. Polyanionic spermine and spermidine, EDTA, Mn²⁺ and Zn²⁺ differentially inhibit the PTPases and can be used to discriminate between phosphatases [28,69-71]. From the first purification of PTPases it was noted that thiol reducing agents must be present during their extraction and isolation to preserve their activity [28,69,72]. This suggested that a cysteine residue was critical for enzyme activity. To pursue this further, sulphhydryl-directed inhibitors were tested for their ability to inhibit PTPase activity. Pot and coworkers showed that three irreversible inhibitors, *N*-ethylmaleimide, *p*-(hydroxymercurio)benzoate and iodoacetate fully inhibited rat LAR's PTPase activity [28]. Equivalent levels of iodoacetamide had no effect on enzyme activity. Analysis of enzyme inactivation at different pHs indicated that the negatively charged iodoacetate interacts with a positive residue at or near the active site of rat LAR, allowing a

reactive cysteine to attack and displace I⁻ from iodoacetate. This was confirmed by experiments using purified rat LAR and ¹⁴C-iodoacetate. Iodoacetate binds rat LAR and the binding is in competition with the substrate. Following the irreversible inactivation of LAR with iodoacetate a stoichiometry of 0.8 mol of iodoacetate bound per mol of enzyme could be demonstrated [73]. A single ¹⁴C-labelled peptide was isolated following proteolytic digestion of LAR, and sequencing of this peptide identified the site of labelling as cysteine 1522, the catalytically essential residue within the central conserved 'HCSAGVGR(S/T)G' PTPase motif. Although there are 11 Cys residues in this protein, only Cys-1522 is labelled, pointing to the unusual reactivity of the active-site thiol residue.

The regulatory effects of phosphorylation on PTPase activity have been examined. CD45 can be readily phosphorylated on serine/threonine by casein kinase 2, protein kinase C, and glycogen synthase kinase 3 *in vitro*, and on tyrosine *in vivo*, but no change in phosphatase activity could be detected [74,75]. Rat LAR is also phosphorylated *in vitro* by protein kinase C and p43^{v-abl} tyrosine kinase but no change in activity was noted with either phosphorylation [28]. The only *in vivo* study showing regulation of a phosphatase by phosphorylation was done by Ostergaard and Trowbridge [76]. They showed that when T-cells were treated with the calcium ionophores ionomycin or A23187, CD45 PTPase activity decreased by 50-90%. Analysis of the CD45 showed that the drop in activity correlated with a drop in serine phosphorylation. Even though the sites or enzymes responsible for this dephosphorylation were not identified, this was the first indication that PTPase activities are regulated by phosphorylation *in vivo*.

The large database of similar PTPase sequences enables conserved residues to be identified and mutated to determine their importance in the catalysis of the enzyme. Systematic mutagenesis of conserved cysteine residues in the catalytic domains of CD45 by Streuli and coworkers indicated that only the first domain catalyzed the dephosphorylation of substrate [15]. Later, a similar conclusion was obtained for human LAR [59]. In this same study other point mutations in the central 'HCSAGVGR(S/T)G' motif were created, identifying the critical importance of this region to catalysis. Pot et al. [28] confirmed that greater than 99% of the activity on the substrates tested arises from the first domain of purified recombinant LAR. Either the second domains of CD45 and LAR act solely in a regulatory manner or appropriate substrates for these domains have not been found. Recently both domains of human PTPase have been shown to contain catalytic activity [71]. Differences in substrate specificity, responses to modulators and the interdependence of these two domains are properties likely to be found important in the action of PTPase *in vivo*.

The ability of a single amino acid mutation to inactivate the PTPase has been a useful negative control for PTPase function, both in vitro and in vivo. Mutation of the cysteine contained in the central PTPase signature motif was used to establish the specificity of PTPase 1 [77], LAR [28], *cdc25* [63,64], the bacterial PTPase in vitro [20], and in vivo [78]. This unique and simple method should be of great utility in future studies of the functional significance of PTPases.

Random chemical mutagenesis has also been utilized to study the stability of PTPases. Tsai and coworkers, using hydroxylamine and *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine, mutated the first cytoplasmic PTPase domain of human LAR and screened for temperature-sensitive and thermostable mutants [79]. This analysis identified 8 temperature-sensitive mutants clustered in an 80-amino-acid region in the first half of the molecule. The grouping of these temperature-sensitive mutants identified a region of this PTPase, highly conserved in other PTPases, that probably is critical for initial folding and stability of the protein. Interestingly, a mutation that causes increased stability and solubility of temperature-sensitive human LAR mutations (C1446-Y) occurs naturally in the rat homologue of LAR and may account for the excellent solubility and stability of the latter enzyme [28].

The studies described above all suggest that the cysteine in the 'HCSAGVGR(S/T)G' motif is an important residue used by PTPases to remove phosphate from phosphoproteins. Kinetic and chemical analysis of the action of PTPases has allowed a probable mechanism of enzyme action to be identified. Guan and Dixon [77] and Pot et al. [28] utilized a rapid denaturation technique to trap a radioactively labelled phospho-cysteine intermediate of rat PTPase 1 and LAR. Formation of this intermediate was transient and dependent on the presence of the catalytically important cysteine. Walsh and his colleagues have also detected a 32 P-labelled phosphoprotein and have eloquently documented an LAR PTPase catalyzed 18 O exchange into phosphate from water [80]. Collectively, these observa-

tions along with work of Zhang and Dixon (unpublished) point to the PTPases utilizing the suggested mechanism outlined in Fig. 3. The nucleophilic cysteine (of the PTPase signature motif) attacks the phosphate of the phosphotyrosine protein substrate, releasing the dephosphorylated protein substrate and forming a thiolphosphate enzyme intermediate. In the second step of the reaction, water attacks this intermediate, regenerating active enzyme and releasing the second product, inorganic phosphate. Kinetic data and 'trapping' of the phospho-enzyme intermediate suggests that the second step in the reaction is rate determining.

IV. Biology of the PTPases

Only a few of the PTPases have a well characterized biological function. Information about PTPase function has been obtained from 'mutant' cells lacking the enzyme. CD45 is expressed on the cell surface of lymphocytes and efforts were directed at the selection of cell lines lacking this surface marker. Because CD45 is critical to T-cell activation, cells lacking functional receptor PTPase have impaired response to T-cell [81,82], CD2 [83], and B-cell receptor stimuli [84]. Intracellular signals induced by CD45 in T-cell activation include phosphoinositol turnover [85] and regulation of cellular serine kinases and phosphatases [76]. Three reviews regarding the biological activity of CD45 have recently appeared [86-88].

Another PTPase whose biological significance is known is the *Yersinia* Yop2B phosphatase. Sequence comparison of the catalytic core of mammalian PTPase showed that the bacterial Yop protein shared a number of invariant residues with the PTPases. Expression of the bacterial protein led to the discovery that it had phosphatase activity [20]. Homologous recombination of the genes encoding catalytically active and inactive enzymes conclusively showed that removal of this PTPase activity renders the bacteria avirulent [78] supporting the important observation that the phosphatase

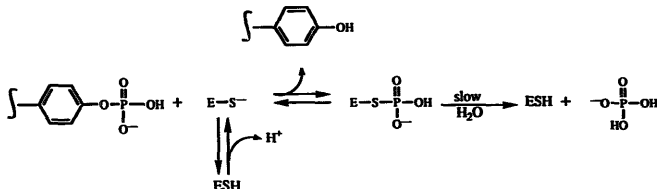


Fig. 3. Suggested mechanism for protein tyrosine phosphatase. E-SH represents the enzyme with the catalytic Cys residue. E-S⁻ denotes the thiol anion of the PTPase which appears to be the reactive form of the enzyme. E-S-PO₃H₂ is the phosphoenzyme intermediate.

activity is essential for virulence in the pathogen which causes diseases such as the plague or the black death.

For some time it had been known that p34^{cdc2}, a serine/threonine kinase contained in the mitosis-promoting factor (MPF) complex of higher eukaryotes, was negatively regulated by tyrosine and serine phosphorylation (reviewed in Ref. 89). Genetic dissection of the system had identified a positive regulatory factor, termed cdc25, that could induce the dephosphorylation of cdc2 and *taus* entry into the cell cycle. Yeast mutants that lack active cdc25 can be induced to enter the cell cycle by the introduction of human T-cell PTPase [90]. Therefore tyrosine phosphatase activity was thought to be contained in or activated by cdc25. Identification of Vh1 [49] gave further impetus to this idea. Alignment of Vh1 to cdc25 (noted by Moreno and Nurse [91]) showed discernible homology, especially in the active-site region of the PTPase. This observation triggered several groups to test the phosphatase activity of purified cdc25. cdc25 has a low-level phosphatase activity toward a very narrow set of substrates [63,64,92,93]. In all cases, p34^{cdc2} was activated by removal of phosphate from its regulatory tyrosine. Parallel research has led to the identification of a PTPase, p65, that has both serine and tyrosine phosphatase activity. This protein is a mitosis-specific component of human MPF [94]. Purified mitotic p65 shows activity towards phosphotyrosine poly(Glu,Tyr) and serine/threonine-phosphorylated histone and phosphor-ylase a [95]. Further evidence supports the identity of this protein as the gene product of cdc25 [95]. The size difference between the predicted 54-kDa human cdc25 [48] and 65-kDa p65 can presumably be accounted for by the latter's *O*-glycosylation. Conclusive identification awaits testing of both proteins with specific antibodies.

During the search for a PTPase responsible for activating p34^{cdc2}, two PTPase homologs from *S. pombe* and *S. cerevisiae* were cloned and characterized [31,32]. Tyrosine phosphatase activity was demonstrated for one of these PTPases [32]. Recently, Guan et al. [96] has reported the cloning of another *S. cerevisiae* PTPase. Ota and Varshovsky [97] have also shown that this *S. cerevisiae* PTPase plays a role in the ubiquitin pathways that are important in protein degradation and in DNA repair.

Because of the growth activating and oncogenic functions of tyrosine kinases, the action of protein tyrosine phosphatases functioning as tumor suppressor genes has also been considered. LaForgia and coworkers pursued this possibility by determining if there was a correlation between deletion of the gene for human PTP γ and the presence of tumors in humans [98]. Results indicated that one allele of the gene was lost in 3 of 5 renal carcinoma cell lines and 5 of 10 lung carcinoma lines. Other PTPase genes have been

mapped to regions of chromosomes with frequent abnormalities associated with human cancer [35,40].

Another method used to identify potential functions of PTPases is to introduce these catalysts into cells and then to evaluate the biological effects of the PTPases. Biological effects observed using this approach include blockage of insulin-like growth factor I receptor autophosphorylation in *Xenopus* oocytes [99], induction of meiotic cell division in *Xenopus* oocytes [100], reduction in baby hamster kidney cell growth rate [101], stimulation of Ca²⁺-dependent amylase secretion from pancreatic acini [102], reduction of transcriptional activation endowed by AP-1 or CREB transcription factors [27], and a reduction in the number of v-src [103] and p185^{neu} [104] transformed foci in NIH-3T3 cells. The exact position where each of these PTPases acts to produce these effects is unknown.

Expression of the PTPases in cells have highlighted aspects of their cellular distributions. Cool and coworkers noted that most of the T-cell PTPase overexpressed in baby hamster kidney cells was associated with the insoluble membrane fraction [101]. Similar results were seen when the same PTPase was expressed in baculovirus [105], where Triton X-100/KCl extractions were necessary to remove the enzyme activity from the cell pellet. Both results suggest a localization of T-cell PTPase to the membrane fraction, possibly through the hydrophobic amino acids at its C-terminus (see above). Results with human PTP1B are similar. Full length forms of the enzyme could only be isolated when the membranes were solubilized in 2% Triton X-100 [106]. When PTP1 was introduced into NIH 3T3 cells the protein was shown to localize to the endoplasmic reticulum and solubility studies indicate that PTP1 had properties of an integral and not a peripheral membrane protein [103]. Recent observations by Neal and his colleagues indicate that the last 35 residues of PTP1B are responsible for this localization to the membrane [107]. Other studies have indicated novel cellular distributions of PTPases. These include the shift of different sized PTPase activities from the particulate to soluble fractions in megakaryocytic differentiating K562 cells [108], a redistribution of CD45 from the Golgi to an unidentified intracellular location upon T-cell activation [109], and the apparent movement of kidney cell PTPase activity from membranes to the cytosol upon cell confluency, with a concomitant decrease in molecular weight and increase in activity [110]. A general activation of low-molecular-weight membrane tyrosine phosphatase activity has also been described by Pallen and Tong in density-dependent growth-arrested fibroblasts [111].

V. Concluding remarks

Knowledge of the protein tyrosine phosphatases has increased dramatically over the last year with the

cloning of new cDNAs, the analysis of recombinant proteins' actions on substrates *in vitro*, deletion analysis of known PTPases *in vivo* to determine biological function, and the evaluation of tyrosine phosphatase effects when introduced into a complex cellular system. It is a diverse and growing family with complex and not yet fully understood functions. It would appear that the proteins will have to be highly regulated, since most likely multiple PTPases will most likely occur in the same cell. The availability of cloned genes, pure proteins, and antibodies to these proteins will give future researchers the tools necessary to continue exploration of this interesting and complex family of regulatory proteins.

Acknowledgements

The authors thank the members of Dr. Jack Dixon's laboratory for critical reading of the manuscript. Computer analysis was performed at the Purdue University ACLCB computing center. This work was supported in part by grants from the National Institutes of Health and the Walther Cancer Institute. David Pot was a Joe Dawson pre-doctoral fellow of the Walther Cancer Institute.

References

- Roach, P.J. (1991) *J. Biol. Chem.* 266 14139-14142.
- Hunter, T. and Sefton, B.M. (1980) *Proc. Natl. Acad. Sci. USA* 77, 1311-1315.
- Collett, M.S., Purchio, A.F. and Erikson, R.L. (1980) *Nature* 285, 167-169.
- Levinson, A.D., Oppermann, H., Varmus, H.E. and Bishop, J.M. (1980) *J. Biol. Chem.* 255, 11973-11980.
- Sefton, B.M., Hunter, T., Ball, E.H. and Singer, S.J. (1981) *Cell* 24, 165-174.
- Bishop, J.M. (1985) *Trends Genet.* 1, 245-250.
- Ushiro, H. and Cohen, S. (1980) *J. Biol. Chem.* 255, 8363-8365.
- Hunter, T. (1987) *Cell* 50, 823-829.
- Fischer, E.H., Charbonneau, H. and Tonks, N.K. (1991) *Science* 253, 401-406.
- Brautigan, D.L. (1992) *Arch. Biochem. Biophys.* in press.
- Tonks, N.K., Diltz, C.D. and Fischer, E.H. (1988) *J. Biol. Chem.* 263, 6722-6730.
- Charbonneau, H., Tonks, N.K., Walsh, K.A. and Fischer, E.H. (1988) *Proc. Natl. Acad. Sci. USA* 85, 7182-7186.
- Thomas, M.L. (1989) *Ann. Rev. Biochem.* 58, 319-369.
- Cool, D.E., Tonks, N.K., Charbonneau, H., Walsh, K.A., Fischer, E.H. and Krebs, E.G. (1989) *Proc. Natl. Acad. Sci. USA* 86, 5257-5261.
- Streuli, M., Krueger, N.X., Tsai, A.Y.M. and Saito, H. (1989) *Proc. Natl. Acad. Sci. USA* 86, 8698-8702.
- Guan, K.-L., Haun, R.S., Watson, S.J., Geahlen, R.L. and Dixon, J.E. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1501-1505.
- Brown-Shimer, S., Johnson, K.A., Lawrence, J.B., Johnson, C., Bruskin, A., Green, N.R. and Hill, D.E. (1990) *Proc. Natl. Acad. Sci. USA* 87, 5148-5152.
- Matthews, R.J., Cahir, E.D., and Thomas, M.L. (1990) *Proc. Natl. Acad. Sci. USA* 87, 4444-4448.
- Sap, J., D'Eustachio, P., Givol, D. and Schlessinger, J. (1990) *Proc. Natl. Acad. Sci. USA* 87, 6112-6116.
- Guan, K.-L. and Dixon, J.E. (1990) *Science* 249, 553-556.
- Kaplan, R., Morse, B., Huebner, K., Croce, C., Howk, R., Ravera, M., Ricca, G., Jaye, M. and Schlessinger, J. (1990) *Proc. Natl. Acad. Sci. USA* 87, 7000-7004.
- Jirik, F.R., Janzen, N.M., Melhado, I.G. and Harder, K.W. (1990) *FEBS Lett.* 273, 239-242.
- Krueger, N.X., Streuli, M. and Saito, H. (1990) *EMBO J.* 9, 3241-3252.
- Lombroso, P.J., Murojoch, G. and Lerner, M. (1991) *Proc. Natl. Acad. Sci. USA* 88, 7242-7246.
- Shen, S.-H., Bastien, L., Posner, B.I. and Chretien, P. (1991) *Nature* 352, 736-739.
- Swarup, G., Kamatkar, S., Radha, V. and Rema, V. (1991) *FEBS Lett.* 280, 65-69.
- Champion-Arnaud, P., Gesnel, M.-C., Foulkes, N., Ronsin, C., Sassone-Corsi, F. and Breathnach, R. (1991) *Oncogene* 6, 1203-1209.
- Pot, D.A., Woodford, T.A., Remboutsika, E., Haun, R.S. and Dixon, J.E. (1991) *J. Biol. Chem.* 266, 19688-19696.
- Chernoff, J., Schievella, A.R., Jost, C.A., Erikson, R.L. and Neel, B.G. (1990) *Proc. Natl. Acad. Sci. USA* 87, 2735-2739.
- Ohagi, S., Nishi, M. and Steiner, D.F. (1990) *Nucleic Acids Res.* 18, 7159.
- Ottlie, S., Chernoff, J., Hannig, G., Hoffman, C.S. and Erikson, R.L. (1991) *Proc. Natl. Acad. Sci. USA* 88, 3455-3459.
- Guan, K.-L., Deschenes, R.J., Qiu, H. and Dixon, J.E. (1991) *J. Biol. Chem.* 266, 12964-12970.
- Gu, M.-X., York, J.D., Warsawsky, I. and Majerus, P.W. (1991) *Proc. Natl. Acad. Sci. USA* 88, 5867-5871.
- Yang, Q. and Tonks, N.K. (1991) *Proc. Natl. Acad. Sci. USA* 88, 5949-5953.
- Gebbink, M.F.B.G., van Etten, I., Hateboer, G., Suijkerbuijk, R., Beijersbergen, R.L., van Kessel, A.G. and Mooleenaar, W.H. (1991) *FEBS Lett.* 290, 123-130.
- Yang, X.-H., Seow, K.T., Bahri, S.M., Oon, S.H. and Chia, W. (1991) *Cell* 67, 661-673.
- Tian, S.-S., Tsoulfas, P. and Zinn, K. (1991) *Cell* 67, 675-685.
- Hariharan, I.K., Chuang, P.-T. and Ruñin, G.M. (1991) *Proc. Natl. Acad. Sci. USA* 88, 11266-11270.
- Mosinger, B. Jr., Tillmann, U., Westphal, H. and Tremblay, M.L. (1992) *Proc. Natl. Acad. Sci. USA* 89, 499-503.
- Yi, T., Cleveland, J.L. and Ihle, J.N. (1992) *Mol. Cell. Biol.* 12, 836-846.
- Plutsky, J., Neel, B.G. and Rosenber, R.D. (1992) *Proc. Natl. Acad. Sci. USA* 89, 1123-1127.
- Zanke, B., Suzuki, H., Kichihara, K., Mizzen, L., Minden, M., Pawson, A. and Mah, T.W. (1992) *Eur. J. Immunol.* 22, 235-239.
- Thomas, M.L., Barclay, A.N., Gagnon, J. and Williams, A.F. (1985) *Cell* 41, 83-93.
- Russell, P. and Nurse, P. (1986) *Cell* 45, 145-153.
- Saga, Y., Tung, J.-S., Shen, F.-W., and Boyse, E.A. (1987) *Proc. Natl. Acad. Sci. USA* 84, 5364-5368.
- Ralph, S.J., Thomas, M.L., Morton, C.C. and Trowbridge, I.S. (1987) *EMBO J.* 6, 1251-1257.
- Streuli, M., Krueger, N.X., Hall, L.R., Schlossman, S.F. and Saito, H. (1988) *J. Exp. Med.* 168, 1523-1530.
- Sadiu, K., Reed, S.I., Richardson, H. and Russell, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 5139-5143.
- Guan, K.-L., Broyles, S.S. and Dixon, J.E. (1991) *Nature* 350, 359-362.
- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olté, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and Venter, J.C. (1991) *Science* 252, 1651-1656.
- Alexander, D.R. (1990) *New Biol.* 2, 1049-1062.

- 52 Feng, D-F. and Doolittle, R.F. (1987) *J. Mol. Evol.* 25, 351-360.
- 53 Edelman, G.M. (1987) *Immunol. Rev.* 100, 11-45.
- 54 Koch, C.A., Anderson, D., Moran, M.F., Ellis, C. and Pawson, T. (1991) *Science* 252, 668-674.
- 55 Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991) *Cell* 64, 281-302.
- 56 Yoshioka, T., Tanaka, O., Otani, H. and Shinohara, H. (1991) *Brain Res.* 555, 177-179.
- 57 Rossmann, M.G., Liljas, A., Branden, C. and Banaszak, L.J. (1975) in *The Enzymes* (Boyer, P.D., ed.), Vol. 11, pp. 61-102. Academic Press, New York.
- 58 Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) *Science* 241, 42-52.
- 59 Streuli, M., Krueger, N.X., Thai, T., Tang, M. and Saito, H. (1990) *EMBO J.* 9, 2399-2407.
- 60 Knighton, D.R., Zheng, J.-H., Ten Eyck, L.F., Ashford, V.A., Xuong, N.-H., Taylor, S.S. and Sowadski, J.M. (1991) *Science* 253, 407-414.
- 61 Hirai, H. and Varmus, H.E. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8592-8596.
- 62 Hirai, H. and Varmus, H.E. (1990) *Genes Dev.* 4, 2342-2352.
- 63 Gautier, J., Solomon, M.J., Booher, R.N., Bazan, J.F. and Kirschner, M.W. (1991) *Cell* 67, 197-211.
- 64 Dunphy, W.G. and Kumagai, A. (1991) *Cell* 67, 189-196.
- 65 Sale, G.J. (1991) *Adv. Prot. Phosph.* 6, 159-186.
- 66 Levine, B.A., Clack, B. and Ellis, L. (1991) *J. Biol. Chem.* 266, 3565-3570.
- 67 Ramachandran, C., Aebersold, R., Tonks, N.K. and Pot, D.A. (1992) *Biochemistry* 31, 4232-4238.
- 68 Cho, H., Ramer, S.E., Itoh, M., Winkler, D.G., Kitas, E., Bannwarth, W., Burn, P., Saito, H. and Walsh, C.T. (1991) *Biochem. J.* 270, 6210-6216.
- 69 Tonks, N.K., Diltz, C.D. and Fischer, E.H. (1988) *J. Biol. Chem.* 263, 6731-6737.
- 70 Mui, L. and Haganir, R.L. (1991) *J. Biol. Chem.* 266, 16063-16072.
- 71 Wang, Y. and Pallen, C.J. (1991) *EMBO J.* 10, 3231-3237.
- 72 Jones, S.W., Erikson, R.L., Ingebritsen, V.M. and Ingebritsen, T.S. (1989) *J. Biol. Chem.* 264, 7747-7753.
- 73 Pot, D.A. and Dixon, J.E. (1992) *J. Biol. Chem.* 267, 140-143.
- 74 Tonks, N.K., Diltz, C.D. and Fischer, E.H. (1990) *J. Biol. Chem.* 265, 10674-10680.
- 75 Stover, D.R., Charbonneau, H., Tonks, N.K. and Walsh, K.A. (1991) *Proc. Natl. Acad. Sci. USA* 88, 7704-7707.
- 76 Ostergaard, H.L. and Trowbridge, I.S. (1991) *Science* 253, 1423-1425.
- 77 Guan, K.-L. and Dixon, J.E. (1991) *J. Biol. Chem.* 266, 17026-17030.
- 78 Bliska, J.B., Guan, K.-L., Dixon, J.E. and Falkow, S. (1991) *Proc. Natl. Acad. Sci. USA* 88, 1187-1191.
- 79 Tsai, A.Y.M., Itoh, M., Streuli, M., Thai, T. and Saito, H. (1991) *J. Biol. Chem.* 266, 10534-10543.
- 80 Cho, H., Ramer, S.E., Itoh, M., Kitas, E., Bannwarth, W., Burn, P., Saito, H. and Walsh, C.T. (1992) *Biochemistry* 31, 133-138.
- 81 Pingel, J.T. and Thomas, M.L. (1989) *Cell* 58, 1055-1065.
- 82 Weaver, C.T., Pingel, J.T., Nelson, J.O. and Thomas, M.L. (1991) *Mol. Cell. Biol.* 11, 4415-4422.
- 83 Koretzky, G.A., Picus, J., Schultz, T. and Weiss, A. (1991) *Proc. Natl. Acad. Sci. USA* 88, 2037-2041, 1991.
- 84 Justement, L.B., Campbell, K.S., Chien, N.C. and Cambier, J.C. (1991) *Science* 252, 1839-1842.
- 85 Koretzky, G.A., Picus, J., Thomas, M.L. and Weiss, A. (1990) *Nature* 346, 66-68.
- 86 Shaw, A. and Thomas, M.L. (1991) *Curr. Opin. Cell Biol.* 3, 862-868.
- 87 Trowbridge, I.S., Ostergaard, H.L. and Johnson, P. (1991) *Biochim. Biophys. Acta* 1095, 46-56.
- 88 Trowbridge, I.S. (1991) *J. Biol. Chem.* 266, 23517-23520.
- 89 Nurse, P. (1990) *Nature* 344, 503-508.
- 90 Gould, K.L., Moreno, S., Tonks, N.K. and Nurse, P. (1990) *Science* 250, 1573-1576.
- 91 Moreno, S. and Nurse, P. (1991) *Nature* 351, 194.
- 92 Starostfeld, U., Labbe, J.C., Fesquet, D., Cavadore, J.C., Picard, A., Sadhu, K., Russell, P. and Doree, M. (1991) *Nature* 351, 242-245.
- 93 Kumagai, A. and Dunphy, W.G. (1991) *Cell* 64, 903-914.
- 94 Meikrantz, W., Suprynovicz, F.A., Halleck, M.S. and Schlegel, R.A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 9600-9604.
- 95 Meikrantz, W., Smith, D.M., Sladicka, M.M. and Schlegel, R.A. (1991) *J. Cell Sci.* 98, 303-307.
- 96 Guan, K.-L., Deschenes, R. and Dixon, J.E. (1992) *J. Biol. Chem.* 267, 10024-10030.
- 97 Ota, I. and Varshovskiy, A. (1992) *Proc. Natl. Acad. Sci. USA* 89, 2355-2359.
- 98 LaForgia, S., Morse, B., Levy, J., Barnea, G., Cannizzaro, L.A., Li, F., Nowell, P.C., Boghosian-Sell, J., Guick, J., Weston, A., Harris, C.C., Drabkin, H., Patterson, D., Croce, C.M., Schlessinger, J. and Huebner, K. (1991) *Proc. Natl. Acad. Sci. USA* 88, 5036-5040.
- 99 Cicirelli, M.F., Tonks, N.K., Diltz, C.D., Weiel, J.E., Fischer, E.H. and Krebs, E.G. (1990) *Proc. Natl. Acad. Sci. USA* 87, 5514-5518.
- 100 Tonks, N.K., Cicirelli, M.F., Diltz, C.D., Krebs, E.G. and Fischer, E.H. (1990) *Mol. Cell. Biol.* 10, 458-463.
- 101 Cool, D.E., Tonks, N.K., Charbonneau, H., Fischer, E.H. and Krebs, E.G. (1990) *Proc. Natl. Acad. Sci. USA* 87, 7280-7284.
- 102 Jena, B.P., Padfield, P.J., Ingebritsen, T.S. and Jamieson, J.D. (1991) *J. Biol. Chem.* 266, 17744-17746.
- 103 Woodford-Thomas, T.A., Rhodes, J.D. and Dixon, J.E. (1992) *J. Cell. Biol.* 117, 401-414.
- 104 Brown-Shiner, S., Johnson, K.A., Hill, D.E. and Bruskin, A.M. (1992) *Cancer Res.* 52, 478-482.
- 105 Zander, N.F., Lorenzen, J.A., Cool, D.E., Tonks, N.K., Daum, G., Krebs, E.G. and Fischer, E.H. (1991) *Biochemistry* 30, 6964-6970.
- 106 Pallen, C.J., Lai, D.S.Y., Chia, H.P., Boulet, I. and Tong, P.H. (1991) *Biochem. J.* 276, 315-323.
- 107 Frangioni, J.V., Beahm, P.H., Shifrin, V., Jost, C.A. and Neel, B.G. (1992) *Cell* 68, 545-560.
- 108 Butler, T.M., Ziemiecki, A. and Friis, R.R. (1990) *Cancer Res.* 50, 6323-6329.
- 109 Minami, Y., Stafford, F.J., Lippincott-Schwartz, J., Yuan, L.C. and Klausner, R.D. (1991) *J. Biol. Chem.* 266, 9222-9226.
- 110 Brautigian, D.L. and Pinault, F.M. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6696-6700.
- 111 Pallen, C.J. and Tong, P.H. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6996-7000.
- 112 Devereux, J., Haeblerli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387-395.