Cloning of murine low molecular weight phosphotyrosine protein phosphatase cDNA: identification of a new isoform

F. Magherini, E. Giannoni, G. Raugei*, P. Cirri, P. Paoli, A. Modesti, G. Camici, G. Ramponi

Dipartimento di Scienze Biochimiche, Università di Firenze, viale Morgagni 50, 50134 Florence, Italy

Received 31 August 1998; received in revised form 14 September 1998

Abstract The low molecular weight phosphotyrosine protein phosphatase (LMW-PTP) is a 18 kDa cytosolic enzyme, involved in the negative regulation of cell proliferation. In different mammalian species LMW-PTPs are expressed in two molecular forms produced from a single primary transcript through an alternative splicing mechanism. In this paper we report the cloning, expression and characterization of mouse isoforms of LMW-PTPs (called m-IF1 and m-IF2), very similar to the corresponding rat and human isoenzymes. Moreover we have identified a third cDNA encoding a protein (m-IF2P) that presents three substitutions compared to m-IF2. This new isoform is still active on pNPP, although to a lower extent: this reduction is mainly due to the leucine to proline substitution in position 13, within the catalytic loop. The mRNA expression level of this isoform is comparable to those of m-IF1 and m-IF2. It is likely that a gene duplication process followed by mutations has generated this new gene.

Key words: Phosphotyrosine protein phosphatase; Alternative splicing; Mouse

1. Introduction

Protein phosphorylation on tyrosine residues is one of the main eukaryotic cell signalling mechanisms and plays a crucial role in cell proliferation and differentiation [1]. Protein tyrosine phosphorylation is transient and is reversibly regulated by the antagonist effects of protein tyrosine kinases (PTKs) and protein tyrosine phosphatases (PTPs) [2]. The PTP family includes more than 70 enzymes, different in molecular weight and subcellular localization.

The low molecular weight protein tyrosine phosphatase (LMW-PTP) subfamily is composed of 18 kDa cytosolic enzymes showing a three-dimensionally folded phosphate loop [3]. LMW-PTPs present the CXXXXXR active site motif, common to all PTPases [4]. It has been demonstrated that both cysteine and arginine are essential for catalysis [2,4]. Since there are no other sequence homologies between LMW-PTPs and the other PTPs, it is very likely that the active site structure common to all PTPs may have originated through convergent evolution of different genes [5,6]. In mammalians (as far as we know) LMW-PTPs are expressed in two molecular forms, termed IF1 and IF2, differing only in the sequence of residues 40–73 [7] and showing different substrate specificity and distinct sensitivity to activating modulators: type 1 (IF1) includes human PTP fast and rat ACP1, while type 2 (IF2) includes human PTP slow, rat ACP2 and bovine LMW-PTP [8]. In man it has been demonstrated that the two isoforms are the product of a single gene through an alternative splicing mechanism [9]. Recently a further human splicing variant (SV3) of LMW-PTP has been characterized, originating from alternative splicing of the same primary transcript [10].

It has been reported that the overexpression of the LMW-PTP in NIH3T3 cells causes a reduction of cell proliferation [11]; this fact indicates that the enzyme plays a role in the control of cell proliferation. In particular LMW-PTP is involved in the regulation of PDGF-dependent mitogenesis [12]. Recently we have demonstrated that LMW-PTP is involved in the regulation of Myc expression dependent on Src activation, and in the regulation of Fos through an ERK-independent pathway mediated by the STAT proteins [13].

In this paper we report the identification of two cDNAs encoding the IF1 and IF2 mouse LMW-PTPs and the isolation of a third cDNA species encoding a new isoform which presents an amino acid substitution in the catalytic loop.

2. Materials and methods

2.1. Isolation of mouse LMW-PTPase cDNAs

The search for DNA sequences encoding mouse LMW-PTP was performed using the human LMW-PTPs as the query sequence in the GenBank and dbEST databases with the BLAST program. Total RNA was purified by the guanidinium isothiocyanate method [14]. First strand cDNA was synthesized using random hexamer primers. DNA amplification was performed by PCR, using two primers (Dir/BamHI 5′-CGCGTGGATCCGCAGAGGTTGGGTCCAAG and Rev/EcoRI 5′-GCGCTGAATTCCTGATGAGTGAGCTGCTC) designed on the basis of 5′ and 3′ regions of the mouse LMW-PTP cDNAs identified by the accession numbers W98119 and AA00119 respectively. PCR products were cloned into PCR II tailing vector using the TA cloning kit (Invitrogen). Recombinant clones were screened with ^{32}P-labeled cDNA probes encoding human LMW-PTP, under low stringency conditions (4×SET, 0.1% SDS and 5×Denhardt's solution, at 40°C overnight). Positive clones were sequenced by the dideoxy chain termination method using Sequenase II polymerase (USB). For the RACE technique, performed according to Frohman and al. [15], primers 2R (5′-CTGTGGATCATAGCATAGCTCCCAAG) and 1R (5′-CCATACATAGTATATAATCG) were used for the first amplification and for the nested PCR respectively.

2.2. Northern blot analysis

Northern blot analysis on total RNA was performed as already

*Corresponding author. Fax: (39) (55) 4222725.
E-mail: raugei@scibio.unifi.it

Abbreviations: LMW-PTP, low molecular weight phosphotyrosine protein phosphatase; PTPs, protein tyrosine phosphatases; PTKs, protein tyrosine kinases; ERK, extracellular signal-regulated kinase; STAT, signal transducer and activators of transcription; pNPP, *para*-nitrophenyl phosphate; RACE, rapid amplification of cDNA ends

PII: S0014-5793(98)01241-1

described [13]. Filters were subjected to autoradiography at −70°C with Kodak films and intensifying screens. Films with appropriate exposure times (in order to have non-saturated signals) were scanned in a Cybertech Image Analyzer for quantitation. The values were normalized to the signals obtained by reprobing the same filters with actin cDNA.

2.3. Determination of LMW-PTP isoform mRNA relative amounts

The RT-PCR products obtained using the primers Dir/BamH1 and Rev/EcoRI (see Section 2.1) were separated in an agarose gel, blotted on nylon membrane filters (Hybond-N^+) and separately hybridized to ^{32}P-labeled oligonucleotides, specific for each cDNA species, namely 1F-Dir (5′-CGGCTACATCCACCTATG) for m-IF1, 1S-Rev (5′-GCAGCTGACAGCTCTTGG) for m-IF2 and Pro13 (5′-GTTCGTGTGTCCTGGTAACAT) for m-IF1P or m-IF2P. Each hybridization was carried out in standard conditions, 2°C below the specific T_m. The efficiency of each hybridization was checked using equal amounts of specific cDNAs spotted on nitrocellulose. The probed nitrocellulose filters were exposed to X-ray films and the positive bands were counted in a β-counter liquid scintillator for quantification. In the attempt to evaluate the level of expression of the m-IF1P form, a RT-PCR experiment specific for the amplification of the m-IF1 isoforms was performed using the IF1-specific fast-Rev primer (5′-CATAGGTGGATGTAGCCGCAC) together with primer Dir/BamHI. The PCR products, after separation on agarose gel, were first hybridized using primer Pro13 as a probe; the same PCR products were then cloned in the PCR II tailing vector (see Section 2.1). The resulting colonies (about 100) were hybridized either with primer Pro13 or with primer 1F-Dir as control. In order to evaluate the concomitant presence of m-IF and m-IFP transcripts in mouse liver, primers Leu13 (5′-GTTCGTGTGTCTCGGTAACAT) and Pro13 were used.

2.4. Site-directed mutagenesis

The Pro^{13}Leu mutant was obtained using the USE mutagenesis kit (Pharmacia), according to Deng and Nickoloff [16]. Mutation was confirmed by DNA sequencing.

2.5. Overexpression, purification and kinetics assays of the proteins

All cDNAs were cloned in frame with glutathione *S*-transferase in the pGEX-2T bacterial expression vector. Recombinant proteins were purified as previously described [17] and analyzed by SDS-polyacrylamide gel electrophoresis [18]. The activities on *p*-nitrophenyl phosphate (pNPP) and the activation of cGMP were determined as previously described [17].

3. Results and discussion

3.1. Isolation and characterization of LMW-PTP cDNAs

Sequence homology analysis on mouse Expressed Sequence Tags databases (dbEST) using the human LMW-PTP cDNA sequence resulted in the individuation of several clones showing a very high degree of identity (more than 90% of the deduced amino acid sequence), very likely representing partial cDNA sequences encoding the mouse LMW-PTP. On the basis of the 5′ (accession number W98119) and 3′ regions (accession number AA00199) of the putative murine LMW-PTP cDNA sequence, two primers (Dir/BamH1 and Rev/EcoR1) were synthesized and used in RT-PCR amplification in order to obtain the complete cDNA. Three cDNA sequences encoding different murine LMW-PTPs (which we named m-IF1, m-IF2 and m-IF2P respectively) were obtained. The deduced amino acid sequences of the three cDNAs are presented in Fig. 1 compared with other mammalian isoforms: m-IF1 and m-IF2 clearly represent the mouse counterparts of the IF1 and IF2 isoforms already characterized in other mammalians, while m-IF2P could represent a new isoform (see below). m-IF1 and m-IF2 most likely originate from an alternative splicing process, in analogy with what has been already demonstrated for the human gene [8]: the two sequences, in fact, present a diverging amino acid sequence in residues 40–72 and complete identity in the rest of the sequence. Moreover, 260 bases of the 5′ end of the cDNA, upstream of the AUG start codon, were obtained with the RACE-PCR technique using mouse NIH-3T3 fibroblast total RNA. This sequence (m-IF2, accession number Y17344), together with m-IF1 (accession number Y17345) and m-IF2P (accession number Y17343) were submitted to the EMBL Nucleotide Sequence Database. Human and murine cDNA show a very

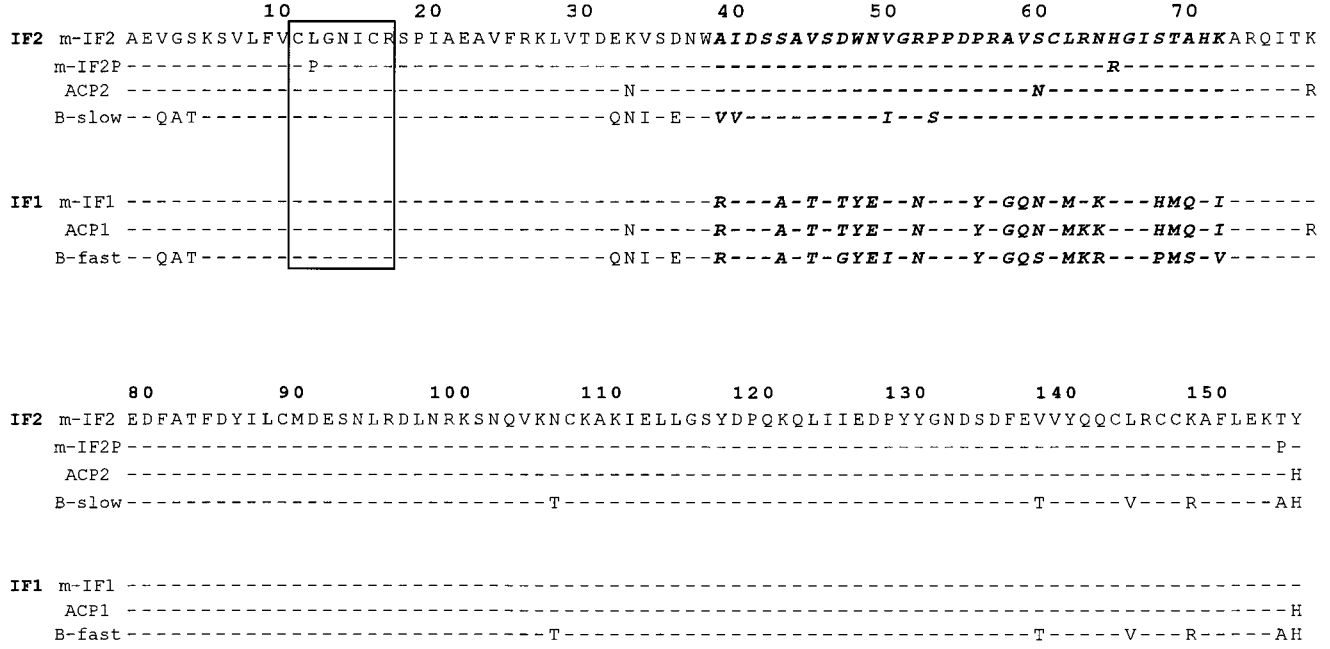

Fig. 1. LMW-PTP amino acid sequence alignment of m-IF2 against m-IF2P, m-IF1 and other mammalian isoforms. ACP1 and ACP2 are from rat, B-fast and B-slow are from man. The amino acids of the catalytic loop are boxed, those of the variable region spanning residues 40–73 are in italics.

high degree of identity (85% and 87% for IF1 and IF2 cDNAs respectively) in the coding sequences. The degree of identity drops to about 36% when the two 5′ untranslated regions are compared. Deduced murine amino acid sequences were compared to the human isoforms: the degree of identity is very high (90.5% and 87.3% for IF1 and IF2 respectively). Remarkably, all the residues known to be important for the catalytic activity, namely Cys^{12}, Arg^{18} and Asp^{129} ([19,20]), are conserved in both sequences. In addition, residues Tyr^{131} and Tyr^{132}, which are known to play a role in the phosphorylation-dependent activation of the enzyme [21], are also present in the murine sequences.

Unexpectedly a third cDNA species was isolated, encoding an isoform presenting three base substitutions compared to the m-IF2 cDNA sequence, leading to three amino acid substitutions in the protein; this isoform, which we have named m-IF2P, presents a proline in position 13 instead of leucine, an arginine in position 66 instead of histidine and a proline in position 156 instead of threonine (Fig. 1). In order to verify whether this cDNA was peculiar to the NIH3T3 cells or whether it may be due to a PCR artefact, RT-PCR was also performed on mouse liver total RNA; this experiment confirmed the presence of the m-IF2P isoform also in the animal tissue.

3.2. Northern blot analysis

Northern blot analysis was performed on total RNA from different adult mouse tissues and from NIH-3T3 fibroblasts. A mixture of m-IF1 and m-IF2 cDNAs was used as probe; with these experiments a single transcript of about 1500 bases in length was always identified. The relative amounts of LMW-PTP mRNA were also evaluated; the results, shown in Fig. 2, demonstrate that LMW-PTP is always expressed, although at very different levels, with a minimum in muscular tissues and a maximum in brain liver.

3.3. Determination of relative amounts of LMW-PTP isoform mRNA levels

In order to evaluate the relative amounts of the different LMW-PTP isoform mRNAs, RT-PCR was performed on total RNA from NIH3T3 cells and from mouse liver, using primers Dir/BamH1 and Rev/EcoR1 for the amplification.

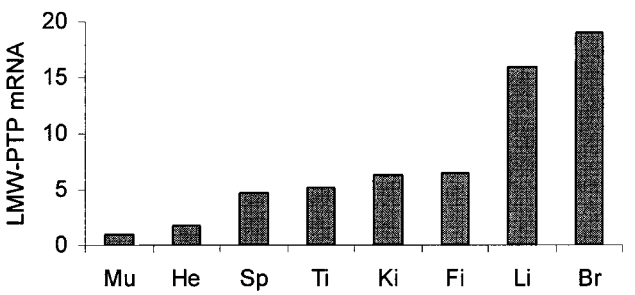

Fig. 2. Relative amounts of murine LMW-PTP in different mouse tissues. The level of mRNA is expressed in arbitrary units, with the value referred to muscle equal to 1. A mixture of m-IF1 and m-IF2 cDNAs was used as a probe. For quantification filters were subjected to autoradiography at −70°C with Kodak films and intensifying screens. Films were scanned in a Cybertech Image Analyser and the values were normalized to the signals obtained by reprobing the same filters with actin cDNA. Mu: skeletel muscle; He: heart; Sp: spleen; Ti: thymus; Ki: kidney; Fi: NIH-3T3 mouse fibroblasts; Li: liver; Br: brain.

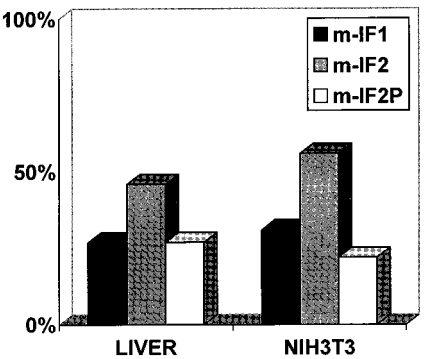

Fig. 3. Determination of the specific mRNA level of each LMW-PTP isoform. Southern blot analysis was performed on the RT-PCR products obtained with primers complementary to sequences common to all the isoforms. Hybridization was performed using, in separate experiments, oligonucleotides specific for each isoform as probes. Quantitation was obtained by counting nitrocellulose bands in a β-scintillator. Quantities are expressed as percentages of the total LMW-PTP transcript.

Since the sequences of these two primers are common to all the LMW-PTP cDNAs previously isolated, the amplification of each sequence should be proportional to the relative amount of mRNA present in the cells. The PCR products were then separately hybridized to ^{32}P-labeled oligonucleotides specific for each isoform cDNA. The results are shown in Fig. 3. The relative mRNA percentages in liver were about 25% for m-IF1, 50% for m-IF2 and 25% for m-IF2P, with very little difference with respect to NIH-3T3 cells. We have to underline that making these calculations is possible only on the basis of another observation: theoretically, as a result of an alternative splicing mechanism, it was reasonable to expect also the existence of the m-IF1P isoform, in addition to the m-IF2P one that we have isolated (it should be noticed that the m-IF1P isoform, if it exists, should contribute to the positive signal obtained using the Pro13 primer as a probe in the previous experiment). In order to assess this point, an RT-PCR experiment was performed on total RNA from NIH3T3 cells and mouse liver with primers BamHI/Dir, complementary to the 5′ end of the murine LMW-PTP mRNA, and fast-Rev, specific for IF1; in this way only the IF1 mRNA species should be amplified. Hybridization of the PCR products with primer Pro13 (specific for the presence of a proline codon in position 13) gave a positive signal, indicating the presence of the m-IF1P mRNA species. Attempts to clone this particular sequence gave negative results: among about 100 m-IF1 clones obtained from either NIH3T3 or liver, no m-IF1P cDNA was isolated. On these basis we conclude that the level of the m-IF1P isoform should be very low in comparison to the other isoforms, making possible the calculation we report above. The fact that m-IF1P is very rare compared to the m-IF2P isoform leads to the hypothesis that additional mutations in this gene (very likely in the introns) could affect the mechanism of splicing, so that the IF1 mRNA is much less abundant than the IF2 one.

The m-IF2P isoform (and possibly also the alternative splicing product m-IF1P) very likely represents the product of a new gene probably derived from duplication of the gene encoding m-IF1 and m-IF2 LMW-PTP. These isoforms could either be the expression of an allelic variant of the m-IF2 gene (as it has been observed in man, where three allelic variants of

Table 1
Kinetic constants and cGMP dependent activation of mouse LMW-PTP isoforms and mutant m-IF2P13L. The parameters were determined using pNPP as substrate and P_i as competitive inhibitor. cGMP activation is expressed as V_{max} fold increase with respect to the negative control.

	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})	K_i (mM)	Maximum cGMP activation
m-IF1	12.3	0.08	154	0.49	1.3
m-IF2	27.3	0.19	143	0.43	9.0
m-IF2P	3.2	15.5	20.8	21.7	1.1
m-IF2P13L	19.3	0.20	96.5	0.29	9.0

LMW-PTP exist) or represent the product of a separate locus. In order to obtain a preliminary indication about this question we performed RT-PCR with the primers Dir/BamH1 and Rev/EcoR1 on total RNA from livers of six outbred mice. Each PCR was separately hybridized with primers Pro13 and Leu13, specific for m-IFP and m-IF forms respectively, in the presence of positive and negative controls. In all cases we obtained a positive signal with both primers, confirming the presence of the two different cDNAs. Since the probability that all six outbred mice were heterozygotes is 1/128, we can hypothesize that the two mRNAs (presenting either proline or leucine in position 13) are the products of two different genes rather than allelic variants of a single gene. In addition the same result was also obtained using total RNA from NIH3T3 cells.

3.4. Expression and characterization of isolated sequences

All isolated sequences were separately expressed in *Escherichia coli*. The main kinetic parameters were measured for the three isoforms and the results are shown in Table 1. The m-IF1 and m-IF2 isoenzymes show kinetic parameter values on pNPP very similar to those of the IF1 and IF2 isoforms [17]. The differences in cGMP-dependent activation between the two mouse isoforms are also very similar to those between the human IF1 and IF2 forms. These results confirm that the m-IF1 and m-IF2 isoforms represent the mouse counterpart of the isoforms already characterized in other mammalians. In contrast the m-IF2P isoform shows, in comparison to the m-IF2 form, a drastically reduced k_{cat} (lower than 12% compared to m-IF2), and enhanced K_m. In addition it has lost the ability to be activated by cGMP. These kinetic differences are probably due to the presence of a proline residue in position 13 in place of a leucine. This amino acid is within the catalytic loop spanning positions 12–18. In order to verify this hypothesis a mutant of the m-IF2P isoform was produced (m-IF2P13L) presenting a leucine in position 13 instead of proline. This mutant showed catalytic properties very similar to those of the m-IF2 isoform (Table 1). These findings confirm that the reduction of enzymatic activity is mainly due to the leucine to proline substitution. The two other differences in the protein sequence do not show any marked influence on the enzymatic activity.

The peculiarity of the fact that in position 13 of m-IF2P an inactivating mutation took place should also be underlined. This position, even if it represents the second residue of the well characterized CXXXXXR catalytic loop, can locate many different amino acids, as observed in many different members of the PTP family; it cannot be excluded that this mutated protein could play a role in the cell, different from that of the two other fully active isoforms.

Acknowledgements: This work was supported by the Italian Association for Cancer Research (AIRC) and in part by the Ministero della Università e Ricerca Scientifica e Tecnologica (MURST-National Project: Meccanismi biochimici di controllo delle funzioni cellulari) and by the Consiglio Nazionale delle Ricerche (Target Project of Biotechnology and CNR 98.00096.PF31).

References

[1] Van der Geer, P., Hunter, T. and Lindberg, R. (1994) Annu. Rev. Cell Biol. 10, 251–337.
[2] Walton, K.M. and Dixon, J.E. (1993) Annu. Rev. Cell Biol. 62, 101–120.
[3] Su, X.D., Taddei, N., Stefani, M., Ramponi, G. and Nordlund, P. (1994) Nature 370, 575–578.
[4] Stone, R.L. and Dixon, J.E. (1994) J. Biol. Chem. 269, 31323–31327.
[5] Barford, D., Flint, A.J. and Tonks, N.K. (1994) Science 263, 1397–1404.
[6] Stuckey, J.A., Shubert, H.L., Fauman, E.B., Zhang, Z.Y., Dixon, J.E. and Saper, M.A. (1994) Nature 370, 571–575.
[7] Ramponi, G. and Stefani, M. (1997) Biochim. Biophys. Acta 1341, 137–156.
[8] Bryson, G.L., Massa, H., Trask, B.J. and Van Etten, R.L. (1995) Genomics 30, 133–140.
[9] Lazaruk, K.D., Dissing, J. and Sensabaugh, G.F. et al. (1993) Biochem. Biophys. Res. Commun. 196, 440–446.
[10] Modesti, A., Marzocchini, R., Raugei, G., Chiti, F., Sereni, A., Magherini, F. and Ramponi, G. (1998) FEBS Lett. (in press).
[11] Berti, A., Rigacci, S., Raugei, G., Degl'Innocenti, D. and Ramponi, G. (1994) FEBS Lett. 349, 7–12.
[12] Chiarugi, P., Cirri, P., Raugei, G., Camici, G., Dolfi, F., Berti, A. and Ramponi, G. (1995) FEBS Lett. 372, 49–53.
[13] Chiarugi, P., Cirri, P., Marra, F., Raugei, G., Fiaschi, T., Camici, G., Manao, G., Romanelli, R.G. and Ramponi, G. (1998) J. Biol. Chem. 273, 6776–6785.
[14] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[15] Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[16] Deng, W.P. and Nickoloff, J.A. (1992) Anal. Biochem. 200, 81.
[17] Cirri, P., Fiaschi, T., Chiarugi, P., Camici, G., Manao, G., Raugei, G. and Ramponi, G. (1996) J. Biol. Chem. 271, 2604–2607.
[18] Laemmli, U.K. (1970) Nature 227, 680–685.
[19] Cirri, P., Chiarugi, P., Camici, G., Manao, G., Raugei, G., Cappugi, G. and Ramponi, G. (1993) Eur. J. Biochem. 214, 647–657.
[20] Taddei, N., Chiarugi, P., Cirri, P., Fiaschi, T., Stefani, M., Camici, G., Raugei, G. and Ramponi, G. (1994) FEBS Lett. 350, 328–332.
[21] Tailor, P., Gilman, J., Williams, S., Couture, C. and Mustelin, T. (1997) J. Biol. Chem. 272, 5371–5374.