Minireview

From structure to function: possible biological roles of a new widespread protein family binding hydrophobic ligands and displaying a nucleotide binding site

Françoise Schoentgen*, Pierre Jollès

Laboratoire des protéines, CNRS URA 1188, Université Paris 5, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France

Received 30 January 1995; revised version received 28 March 1995

Abstract **A cytosolic 21–23 kDa protein isolated from bovine brain was demonstrated to bind hydrophobic ligands, particularly phosphatidylethanolamine. The protein was encountered in numerous tissues of several species. High expression of the mRNA encoding the 21–23 kDa protein was found in rat testes. Immunohistochemical studies showed the presence of the 21–23 kDa protein in the elongated spermatids and epididymal fluid of rat testis and in brain oligodendrocytes of developing rats. As the bovine, human and rat brain 21–23 kDa proteins had only few sequence homologies with already known proteins, it was concluded that they belong to a new protein family. In order to get additional information on the structural features of the 21–23 kDa protein, we built a molecular model which displayed a nucleotide binding site. The affinity of the bovine brain 21–23 kDa protein towards nucleotides as well as its association with cytosolic proteins and small GTP-binding proteins were demonstrated. Recently, significant sequence homologies were found with an antigen from *Onchocerca volvulus*, a fruit fly odorant-binding protein and the yeast protein TFS1 which is a dosage-dependant suppressor of CDC25 mutations. A positive regulation of RAS is carried out by CDC25 product which facilitates the GDP/GTP exchange on RAS proteins. These results imply that 21–23 kDa proteins function in oxidoreduction reactions and signal mechanisms during cell growth and maturation.**

Key words: 21–23 kDa protein; Hydrophobic ligand; Phosphatidylethanolamine-binding protein; Phospholipid-binding protein

1. Introduction

During the 1980s several soluble cytosolic proteins implicated in the binding of hydrophobic ligands have been discovered and extensively studied. Among these proteins were the sterol carrier proteins [1], the phospholipid exchange proteins [2], the retinol and retinoic acid-binding proteins [3,4], the glycolipid transfer proteins [5] and the fatty acid-binding proteins [6]. While important metabolic functions have been described for these proteins, their biological role is often not yet completely elucidated. We identified a basic, cytosolic protein from bovine brain [7], belonging to this hydrophobic ligand binding-protein family. It had a molecular mass of about 23 kDa and was shown to be an anion-binding protein; later, it was characterized as a phosphatidylethanolamine binding protein [8]. The amino acid sequence showed that the 23 kDa protein might represent a new protein family and that its correct molecular weight was 21 kDa [9]; thus, it was called 21–23 kDa protein. Since then, it was encountered in numerous species and tissues; moreover, based on amino acid sequence similarities, other proteins were identified belonging to the same 21–23 kDa protein family. It seems of interest to review the present knowledge concerning this new protein family in order to get information about its potential biological role(s).

2. The 21–23 kDa protein is present in numerous tissues from several species

By using a specific antibody directed against the bovine brain 21–23 kDa protein, Bernier et al. [7] demonstrated the presence of the protein in bovine liver as well as in soluble extracts of rat and mouse brain and in human platelets [7]. In 1988, Bollengier and Mahler [10] reported the presence of the 21–23 kDa protein in several tissues: brain, liver, heart, spleen, stomach, cardiac and skeletal muscle from human, bovine, rat and chicken species. In all mammalian species investigated, the 21–23 kDa protein appeared to be particularly abundant in brain, muscle and liver. In rat brain, Roussel et al. [11] showed that the anti-21–23 kDa protein antibody produced an intense immunochemical staining of oligodendrocytes in the central nervous system of developing rat. The electron microscopic study revealed that the labeling covers the entire cytoplasm of the oligodendrocytes, being more dense along the membranes of rough endoplasmic reticulum and the plasma membrane. Other cytoplasmic organelles were unstained. More recently, a protein identified to be the 21–23kDa protein was encountered in high amounts in rat and mouse testis and epididymis [12–14]: it was localized essentially in the epididymal fluid and the old spermatids.

3. Structural features of the 21–23 kDa protein from mammalian brain

In 1980, Bollengier and Mahler [15] purified from human brain a cytosolic protein identified as h3 which behaved in SDS-PAGE as a component corresponding to $M_r \approx 23$ kDa. The determination of the 24 amino acid N-terminal sequence of protein h3 revealed 75% identity with the corresponding sequence of the bovine brain 21–23 kDa protein [7]. Recently, the

*Corresponding author. *Present address:* Centre de Biophysique Moléculaire, CNRS UPR 4301, rue Charles Sadron, 45071 Orléans Cedex 02, France.

0014-5793/95/$9.50 © 1995 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(95)00376-2

human protein h3 primary structure was established both by deduction from cDNA sequence and by direct microsequencing of the purified protein [13]: it contained 186 amino acids corresponding to a calculated M_r of 21 kDa and was 95% identical to the bovine 21–23 kDa protein. When studying rat brain opioid receptor by morphine affinity chromatography, Grandy et al. [16] isolated independently a 23 kDa protein from rat brain membranes. The authors screened a rat brain cDNA library and established the sequence of a cDNA encoding the 23 kDa rat brain protein. The deduced 23kDa protein shared 85% identity with the bovine 21–23 kDa protein and 85.5% with the human protein h3, revealing that the three mammalian proteins are members of the same 21–23 kDa protein family, the observed amino acid substitutions being due to species differences [13]. The rat brain 23 kDa protein was able to bind morphine derivatives, but was not an opioid receptor itself; however, it may be associated with such a receptor [16].

When studying the affinity of the bovine brain 21–23 kDa protein for phosphatidylethanolamine, Bernier et al. [8] observed that addition of increasing amounts of phosphatidylethanolamine to a constant amount of protein induced the formation of polymers. With low amounts of phosphatidylethanolamine, up to a molecular ratio (protein/phosphatidylethanolamine) of about 1:4, no polymerization occurred, while with increasing phospholipid concentrations SDS/PAGE demonstrated first the appearance of dimers and later, of polymers. Furthermore, physicochemical studies showed that the human brain 21–23 kDa (protein h3) has two free sulphydryl groups responsible for doublets (21 and 23 kDa) and polymer formation by disulphide bonding [10]. Thus, it appeared that in the presence of oxygen as well as in the presence of ligands, the 21–23 kDa protein is able to polymerize. The cysteine residues are known to be implicated in the polymerization mechanisms due to oxygen; however it is possible that during ligand-binding, other residues, particularly the tyrosine residues (which are able to form bis-Tyr) could participate in dimerization and polymerization mechanisms.

In order to try to find out the real biological role of the 21–23 kDa protein, we tried to get information from the major sequence and 3D structure databanks. Comparison between the bovine brain 21–23 kDa protein primary structure and sequence data banks using Hydrophobic Cluster Analysis (HCA) [17] showed the best similarity with the N-terminal domain of phosphoglycerate kinases. Based on this result, a molecular model of the bovine 21–23 kDa protein was built from the crystallographic data of the *Saccharomyces cerevisiae* phosphoglycerate kinase. This model contained a consensus sequence and three-dimensional folding corresponding to a potential nucleotide binding site [18]. The binding of nucleotides (especially FMN and GTP) to the bovine brain 21–23 kDa protein was then demonstrated by affinity chromatography [19]. The 21–23 kDa protein was shown to associate with cytosolic proteins and small GTP-binding proteins. Our results suggested that through its association to small GTP-binding proteins, the 21–23 kDa protein could be implicated in signal mechanisms during cell growth and maturation [19].

4. The testicular protein homologous to the brain 21–23 kDa protein

The cDNA encoding the human brain 21–23 kDa protein allowed us to test the expression of the corresponding mRNA in different tissues from human, rat and mouse species [13]. Multiple tissue Northern blots revealed the presence of a single mRNA in the different tissues of each species: a single band of 1.8 kb, 1.45 kb and 1.2 kb was detected in human, mouse and rat tissues, respectively. Moreover, the expression rate of the mRNA encoding the 21–23 kDa protein appeared to be variable according to the considered tissues. It is striking that the mRNA is particularly highly expressed in rat and mouse testis. Indeed, in rat testis, the level of mRNA encoding the 21–23 kDa protein was 30 times higher than the mRNA encoding the rat brain 21–23 kDa protein [13]. On the contrary, the mRNA was not observed in human testis by Northern blot analysis and was only detected by PCR amplification [13].

These results prompted us to study by immunohistochemistry the location of the 21–23 kDa protein in rat and mouse testis. The protein was immunodetected in high amounts in late spermatids; moreover, our studies showed the labeling of residual bodies, cytoplasmic droplets and to a lesser extent of Sertoli cells suggesting that the protein is essentially expressed in elongated spermatids, concentrated in cytoplasmic droplets then finally resorbed by Sertoli cells [20]. Several authors previously observed in different parts of the rat male reproductive organs a protein considered now to be the 21–23 kDa protein despite some discrepancy concerning the molecular weight evaluation. Indeed, Jones and Hall [21] have purified a 23 kDa protein identified as a major component of rat epididymal secretions and sperm plasma membranes; according to these authors the 23 kDa protein encountered in rat testis was identical to a 22 kDa protein previously described by Brooks [22]. More recently, in studying the sequential interactions of epididymal secretory proteins with spermatozoa during epididymal transit, a protein of 25 kDa was observed which remained associated with spermatozoa in substantial amounts during epididymal transit [12]. It appeared to be synthesized, secreted and bound to spermatozoa in the proximal epididymis. Partial sequence analyses suggested that the rat testicular 25 kDa protein was identical to the rat brain 21–23 kDa protein [13,21,23]. Recently, these results were corroborated by the establishment of the complete amino acid sequence of the rat epididymal protein; indeed, Perry et al. [14] cloned and sequenced a cDNA encoding the rat epididymal protein which was identical to the rat brain protein [16]. The complete amino acid sequence of the monkey epididymal protein was also determined [14]; only five amino acid substitutions where found compared to the human brain 21–23 kDa protein. On SDS-PAGE, the rat testicular protein migrated with a higher apparent M_r (about 25 kDa) than the protein present in mammalian brain (about 23 kDa). At present, there is no explanation for this discrepancy: it might be due to differences in post-translational events or to 3D structure modifications. However, according to its apparent higher molecular weight, we call the protein encountered in the testis 'testicular phosphatidylethanolamine binding protein' (tPBP) in reference to the initial properties shown for the bovine brain 21–23 kDa protein [8] and to the evidence that the testicular protein is able to bind phosphatidylethanolamine in vitro [21].

Since 1985, by using antibodies directed against tPBP, Brooks [22] observed that the protein could not be detected in blood, peritoneal fluid, saliva, milk, uterine fluid, seminal vesicle secretion, coagulating gland secretion or prostatic secretion.

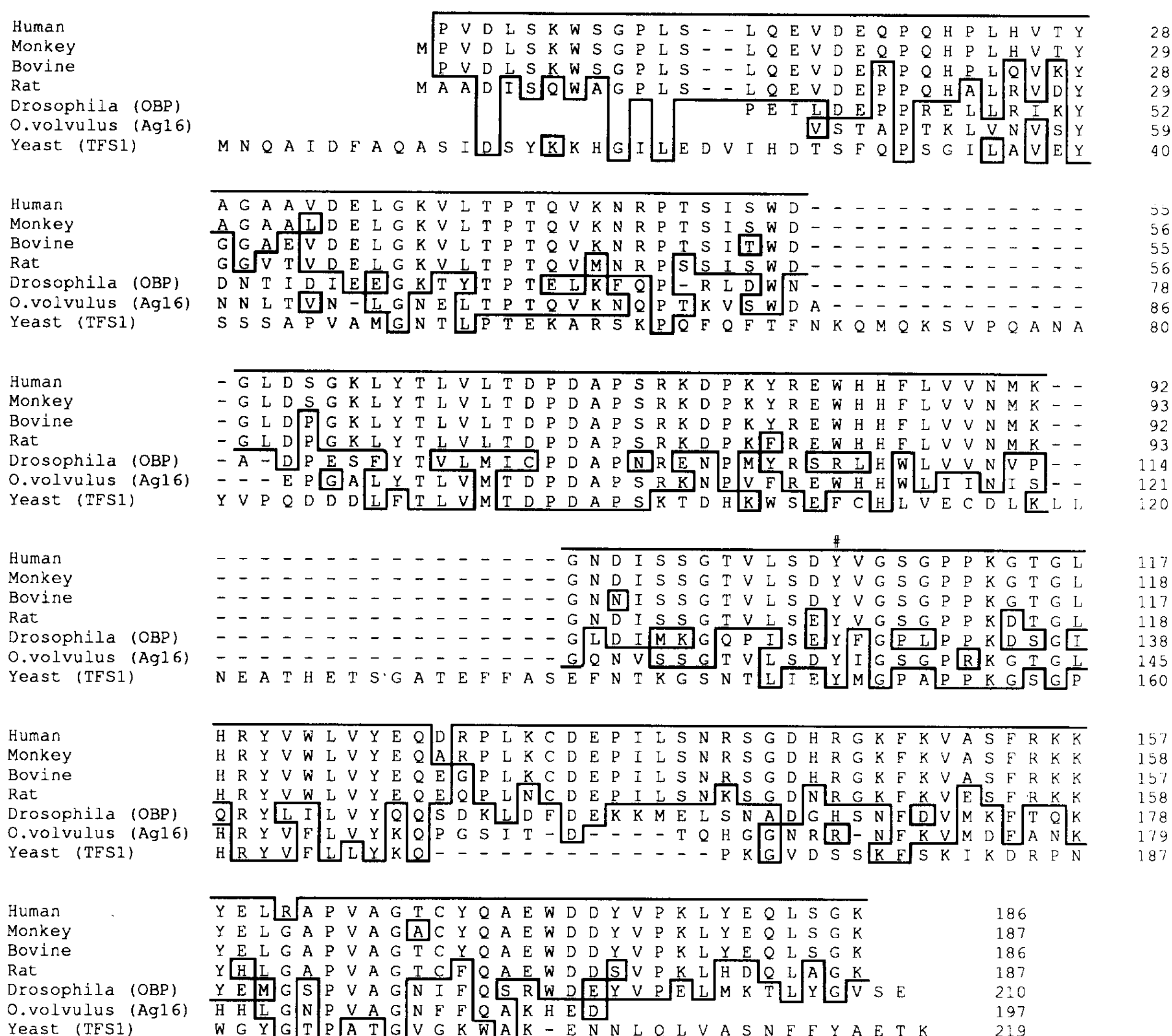

Fig. 1. Comparison between the 21–23 kDa protein family members. Human brain 21–23 kDa protein [13], monkey (*Macaca fascicularis*) epididymal tPBP [14], bovine brain 21–23 kDa protein [9], rat brain 21–23 kDa protein [16], *Drosophila* putative odorant-binding protein [26], *Onchocerca volvulus* Ag16 [24], *Saccharomyces cerevisiae* TFS1 product [25]. Identical amino acids are boxed. # = the C-terminal part of *O. volvulus* antigen was deduced by translation of the second reading frame of the published cDNA sequence [24].

Consequently, it seemed that the testis and epididymis may be unique in containing the protein in a soluble form within their luminal secretions. In addition to the occurrence of the tPBP as a soluble moiety in rat testicular and epididymal fluids, the protein was also located on sperm plasma membranes where its distribution was restricted to the surface of the flagellum. Later on, when studying interactions of epididymal secretory proteins with spermatozoa during epididymal transit, Cornwall et al. [12] observed that the tPBP appeared to be synthesized, secreted and bound to spermatozoa in the proximal epididymis, the protein remaining bound to cauda spermatozoa after 7 days. Further studies [23] showed that the tPBP was present in the luminal fluid of the epididymis that had been separated from the testis prior to the onset of spermatogenesis, demonstrating that the tPBP is present independently in the testis and the epididymis. Recently, Perry et al. [14] disagreed with this point of view: considering the absence in tPBP of a classical N-terminal signal peptide encountered in secreted proteins, they proposed that the tPBP present in rete testis fluid, instead of being secreted, is released from spermatozoa.

5. Proteins from fly, nematode and yeast are members of the 21–23 kDa protein family

Amino acid sequence similarities were found between the 21–23 kDa proteins from mammals and proteins recently purified from other organisms. Significant sequence identities were observed with an immunodominant antigen (Ag16) of the filarial parasite *Onchocerca volvulus* [24], the *Saccharomyces cerevisiae* TFS1 gene product [25] and a putative odorant-binding protein from *Drosophila melanogaster* [26]. The Ag16 sequence was deduced from cDNA analysis; the open frame encoded a 152 amino acid-long protein which showed 37% identity with the bovine brain 21–23 kDa protein. This identity rate could be increased to 44% by changing the reading frame at position 451 of the cDNA [27] (Fig. 1). The Ag16 was localized in the

hypodermis, the cuticle and the uterine epithelium of the parasite; however, its biological function remains unknown [24]. Between the yeast TFS1 gene product and the bovine brain 21–23 kDa protein 29% of the amino acids are identical and the similarity increased to 53% when conservative changes were included [27]. The TFS1 product was described as a dosage-dependent suppressor of CDC25 mutations. The CDC25 product is known to be implicated in the GDP/GTP exchange at the RAS1 and RAS2 levels. When the RAS proteins are GTP-bound they activate the adenylate cyclase which is implicated in the control of cell growth. CDC25 is implicated in at least two different nutrient-specific signalling pathways in yeast starved cells: in addition to the glucose-induced activation of adenylyl cyclase, it controls a nitrogen-specific signalling pathway involving the effector phosphoinositide-specific phospholipase C [28]. Adenylyl cyclase activity and CDC25 gene product were described to be partitioned between soluble and membrane fractions [29]. The sequence homology observed between TFS1 and the brain 21–23 kDa protein suggested that the latter could be implicated in the control of cerebral cell growth and maturation [27]. The putative odorant-binding protein (OBP) was detected in drosophila hairs [26]. A polymerase chain reaction-based method was used to clone a cDNA encoding this protein and in situ hybridization led to the localization of OBP in scattered sensilla on the posterior surface of the fly antenna. The OBP has a putative signal sequence at its amino terminus suggesting that it is secreted into the lumen of olfactory hairs [26].

6. Potential biological roles of the 21–23 kDa protein family

It is noteworthy that the real biological function(s) of the proteins belonging to the 21–23 kDa protein family remain(s) still to be defined; however, the protein family members display several common features and properties at the molecular and cellular levels.

At the molecular level, the capacity of the 21–23 kDa protein family to bind hydrophobic ligands is confirmed; indeed, the bovine brain protein was shown to bind phosphatidylethanolamine [8], the rat brain protein was purified using its affinity for morphine [16], the rat tPBP was shown to bind phosphatidylethanolamine [21] and the putative OBP in *Drosophila* may be able to bind odorant molecules [26]. Furthermore, when comparing the amino acid sequences of the 21–23 kDa protein family members, the central region constituted by residues 60–126 appeared to be particularly well conserved (see Fig. 1); interestingly, this central region corresponds to the potential nucleotide binding site observed in the bovine protein model [18]. The nucleotide site may be implicated in the binding of the bovine brain 21–23 kDa protein to small GTP-binding proteins [19]; the possibility for the mammalian protein to bind small G proteins is in accordance with the observation that TFS1 seems to be implicated in pathway modulations involving small GTP-binding proteins.

At the cellular level, the fact that the brain 21–23 kDa protein was essentially encountered in oligodendrocytes of developing rats and the tPBP in rat and mouse late spermatids indicate the presence of 21–23 kDa protein family members in late stages of growth and maturation of the cells. Moreover, inside the cells of brain and testis, the partitioning of 21–23 kDa protein and tPBP, respectively, between cytoplasm and plasma membranes reinforce the suggestion of their possible implication in cell growth and maturation. Recently, Onoda and Djakiew [30] have proposed that tPBP could be a germ cell product responsible for paracrine regulation of Sertoli cell function. This hypothesis seems premature because of the ubiquitous localization of the protein in seminiferous tubular cells and considering that immunoprecipitation of the round spermatid protein had little effect on the down regulation of Sertoli cell secretion and was only an indirect evidence. In testis, tPBP was also located outside the cells on sperm membranes where its distribution was restricted to the surface of the flagellum [22]. Among sperm surface proteins, tPBP was the major protein containing SH groups and one of the major entities containing disulphide bonds. According to Brooks [22], these properties may be of importance in the maintenance of sperm viability. Furthermore, it appears that the testicular protein pattern of expression may imply a role in the regulation of sperm maturation, motility and/or fertilization.

The observations made in testis, compared with the results in rat brain where oligodendrocytes are particularly stained in developing rats, suggest that the 21–23 kDa protein family may be implicated in cell growth, more especially in cell elongation. During this process, the protein family members could be associated directly or indirectly with the modulation or modification of membranes structures. Indeed, in brain, the main function of oligodendrocytes is to elaborate myelin sheaths and to maintain their integrity. Moreover, myelin contains a high amount of lipids and more particularly of ethanolamine phosphoglycerides. In testis, it is known that during maturation of spermatozoa in the epididymis there is considerable remodeling of the plasma membrane, especially in the content of ethanolamine phosphoglycerides which decreases approximately 3-fold. The 21–23 kDa protein in brain and its counterpart in epididymal plasma could have a role as a lipid carrier protein participating in membrane constitution and preventing formation of lysophosphatides known to perturb membrane structure [21].

7. Conclusion

The 21–23 kDa protein family appears to constitute a new type of protein both from a structural and a functional point of view. The biological function of this protein family is not completely known; however, pieces of information are available and the results obtained with purified proteins are in accordance with several observations made on living cells. (1) In vitro, the soluble brain 21–23 kDa protein is able to bind hydrophobic ligands as well as physiological membranes; in vivo, the protein is partitioned between cytoplasm and cell membranes. (2) Immobilized on nitrocellulose sheets, the brain 21–23 kDa protein binds small GTP-binding proteins; in the living yeast, the 21–23 kDa protein family member TFS1 suppresses the effect of the mutations of CDC25 which is implicated in the GDP/GTP exchange of RAS1 and RAS2. (3) In the presence of oxygen, the purified 21–23 kDa protein forms dimers and polymers; among sperm surface proteins, tPBP is the major protein containing free SH groups. Taken all together, these observations strongly suggest the involvement of the 21–23 kDa protein family in two main mechanisms: signal transduction, probably through association with small GTP-binding proteins, and oxidoreduction reactions. Such a role would agree

with 21–23 kDa protein expression during membrane remodeling, cell growth and maturation.

References

[1] Noland, B.J., Arebolo, R.E., Hansbury, E. and Scallen, T.J. (1980) J. Biol. Chem. 255, 4282–4289.
[2] Kader, J.C., Douady, D. and Mazliak, P. (1982) in: New Comprehensive Biochemistry (Hawthorne, J.N.and Ansell, G.B. eds.) vol. 4, pp. 279–311, Elsevier, Amsterdam.
[3] Sundelin, J., Anundi, H., Tragardh, L., Eriksson, U., Lind, P., Ronne, H., Peterson, P.A. and Rask, L. (1985) J. Biol. Chem. 260, 6488–6493.
[4] Sundelin, J., Das, S.R., Eriksson, U., Rask, L. and Peterson, P.A. (1985) J. Biol. Chem. 260, 6494–6499.
[5] Abe, A. and Sasaki, T. (1985) J. Biol. Chem. 260, 11231–11239.
[6] Takahashi, K., Odani, S. and Ono, T. (1983) Eur. J. Biochem. 136, 589–601.
[7] Bernier, I. and Jollès, P. (1984) Biochim. Biophys. Acta 790, 174–181.
[8] Bernier, I., Tresca, J.P. and Jollès, P. (1986) Biochim. Biophys. Acta 871, 19–23.
[9] Schoentgen, F., Saccoccio, F., Jollès, J., Bernier, I. and Jollès, P. (1987) Eur. J. Biochem. 166, 333–338.
[10] Bollengier, F. and Mahler, A. (1988) J. Neurochem. 50, 1210–1214.
[11] Roussel, G., Nussbaum, F., Schoentgen, F., Jollès, P. and Nussbaum, J.L. (1988) Dev. Neurosci. 10, 65–74..
[12] Cornwall, G.A., Vreeburg, J.T., Holland, M.K. and Orgebin-Crist, M.C. (1990) Biol. Reprod. 43, 121–129.
[13] Seddiqi, N., Bollengier, F., Alliel, P.M., Périn, J.P., Bonnet, F., Bucquoy, S., Jollès, P. and Schoentgen, F. (1994) J. Mol. Evol. 39, 655–660.
[14] Perry, A.C.F., Hall, L., Bell, A.E. and Jones, R. (1994) Biochem. J. 301, 235–242.
[15] Bollengier, F. and Mahler, A. (1980) Neuropeptides 1, 119–135.
[16] Grandy, D.K., Hanneman, E., Bunzow, J., Shih, M., Machida, C.A., Bidlack, J.M. and Civelli, O. (1990) Mol. Endocrinol. 4, 1370–1376.
[17] Gaboriaud, C., Bissery, V., Benchetrit,T. and Mornon, J.P.(1987) FEBS Lett. 224, 149–155.
[18] Schoentgen, F., Seddiqi, N., Bucquoy, S., Jollès, P., Lemesle-Varloot, L., Provost, K. and Mornon, J.P. (1992) Protein Eng. 5, 295–303.
[19] Bucquoy, S., Jollès, P. and Schoentgen, F. (1994) Eur. J. Biochem. 225, 1203–1210
[20] Seddiqi, N., Segretain, D., Bucquoy, S., Pineau, C., Jégou, B., Jollès, P. and Schoentgen, F. (1995) Experientia (in press).
[21] Jones, R. and Hall, L. (1991) Biochim. Biophys. Acta 1080, 78–82.
[22] Brooks, D.E. (1985) Biochim. Biophys. Acta 841, 59–70.
[23] Araki, Y., Vierula, M.E., Rankin, T.L., Tulsiani, D.R.P. and Orgebin-Crist, M.C. (1992) Biol. Reprod. 47, 832–843.
[24] Lobos, E., Altmann, M., Mengod, G, Weiss, N., Rudin, W. and Karam, M. (1990) Mol. Biochem. Parasitol. 39, 135–146.
[25] Robinson, L.C. and Tatchell, K. (1991) Mol. Gen. Genet. 230, 241–250.
[26] Pikielny, C.W., Hasan, G., Rouyer, F. and Rosbash, M. (1994) Neuron 12, 35–49.
[27] Schoentgen, F., Bucquoy, S., Seddiqi, N. and Jollès, P. (1993) Int. J. Biochem. 25, 1699–1704.
[28] Shomerus, C. and Küntzel, H. (1992) FEBS Lett. 307, 249–252.
[29] Jones, S., Vignais, M.L. and Broach, J. (1991) Mol. Cell. Biol. 11, 2641–2646.
[30] Onoda, M. and Djakiew, D.(1993) Biochem. Biophys. Res. Commun. 197, 688–695.