Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*

Jean-Benoit Frenette Charron, Ghislain Breton, Mohamed Badawi, Fathey Sarhan*

Département des Sciences Biologiques, Université du Québec à Montréal, P.O. Box 8888, Succ. ''Centre-Ville'', Montréal, QC, Canada H3C 3P8

Received 25 February 2002; revised 11 March 2002; accepted 11 March 2002

First published online 29 March 2002

Edited by Ulf-Ingo Flügge

Abstract Two cDNAs corresponding to a novel lipocalin were identified from wheat and *Arabidopsis*. The two cDNAs designated *Tatil* for *Triticum aestivum* L. temperature-induced lipocalin and *Attil* for *Arabidopsis thaliana* temperature-induced lipocalin encode polypeptides of 190 and 186 amino acids respectively. Structure analyses indicated the presence of the three structurally conserved regions that characterize lipocalins. Sequence analyses revealed that this novel class of plant lipocalin shares homology with three evolutionarily related lipocalins: the mammalian apolipoprotein D (ApoD), the bacterial lipocalin and the insect Lazarillo. The comparison of the putative tertiary structures of both the human ApoD and the wheat *Ta*TIL suggest that the two proteins differ in membrane attachment and ligand interaction. Northern analyses demonstrated that *Tatil* and *Attil* transcripts are upregulated during cold acclimation and heat-shock treatment. The putative functions of this novel class of plant lipocalins during temperature stresses are discussed. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Wheat; Cold acclimation; Heat-shock; Apolipoprotein D; Lipocalin; Freezing tolerance

1. Introduction

Lipocalins are a large and diverse group of small, mostly extracellular proteins that are found in vertebrates and invertebrate animals, plants and bacteria. They are characterized by a conserved ligand binding pocket, which gives them the ability to bind small, principally hydrophobic molecules. The list of potential ligands is constantly growing and includes diverse molecules such as steroids, pheromones, and odorant molecules. Lipocalins were early established as transport proteins, but it is becoming increasingly clear that some of them may be implicated in many other important functions such as modulation of cell growth and metabolism, binding of cell-surface receptors, nerve growth and regeneration, regulation of the immune response, smell reception, cryptic coloration, membrane biogenesis and repair, and induction of apoptosis [1]. In plants, only one study reported the presence of two lipocalin-like proteins. Those lipocalins were found to be key enzymes of the xanthophyll cycle responsible for the protection against photo-oxidative damage [2].

Considering the wide distribution of lipocalins in animals, insects and bacteria, and their expression under conditions of environmental stress [1], it is important to survey the plant genome for stress-regulated lipocalin-like proteins. Towards this goal, we searched our expressed sequence tags (ESTs) database, generated from cold-acclimated wheat, and that of *Arabidopsis* ESTs collection from GenBank™, to determine if some of the clones isolated so far present sequence homology with lipocalins. Two homologous cDNAs, from wheat (*Tatil*) and *Arabidopsis thaliana* (*Attil*), were found to encode proteins that contain the three structurally conserved regions (SCRs) that characterize lipocalins. Those proteins share significant homology with three established members of the lipocalin family, the mammalian apolipoprotein D (ApoD), the bacterial lipocalin Blc, and the insect Lazarillo. The structure, regulation and putative function of this novel class of plant lipocalin during environmental stresses are discussed.

2. Materials and methods

2.1. Plant material and growth conditions

In this study we used two spring wheat genotypes (*Triticum aestivum* L. cv Glenlea and cv Concorde), four winter wheat genotypes (*T. aestivum* L. cv Monopole, cv Absolvent, cv Fredrick, and cv Norstar), winter rye (*Secale cereale* L. cv Musketeer), oat (*Avena sativa* L. cv Laurent), barley (*Hordeum vulgare* L. cv Winchester) and *A. thaliana* ecotype Columbia.

Cereal plants were germinated in moist sterilized vermiculite for 5 days in the dark and 2 days under artificial light (225 μE) at 25°C/20°C (day/night). Control plants were maintained under the same conditions while cold acclimation and other stress treatments were performed as previously described [3]. *Arabidopsis* plants were grown in pots in a 1:1 mixture of Promix (Premier) and vermiculite in a growth cabinet for 40 days under 8 h artificial light (70 μE) at 22°C (70% relative humidity). For cold treatment, plants were placed under the same light conditions at 4°C. For heat-shock treatment the plants were exposed to 45°C for 1 h. This condition elicits a typical heat-shock response.

2.2. Cloning and molecular analysis

The wheat lipocalin cDNA was isolated from a Lambda Zap II library (Stratagene) constructed from poly(A)$^+$ RNA that was isolated from 1 day cold-acclimated winter wheat (*T. aestivum* L. cv Norstar) [4]. The *Arabidopsis* lipocalin was identified by homology search [5] using the wheat sequence against the *Arabidopsis* ESTs database. The identified clone (ID: 120O12) from the PRL2 library [6] was ordered from the Arabidopsis Biological Resource Center. Complete DNA sequences of wheat and *Arabidopsis* clones were determined from both strands. Cereals RNA extractions were performed

*Corresponding author. Fax: (1)-514-987 4647.
E-mail address: sarhan.fathey@uqam.ca (F. Sarhan).

Abbreviations: ABA, abscisic acid; *At*, *Arabidopsis thaliana*; ApoD, apolipoprotein D; Blc, bacterial lipocalin; GPI, glycosylphosphatidylinositol; SCR, structurally conserved region; SSC, saline sodium citrate; *Ta*, *Triticum aestivum* L.; TIL, temperature-induced lipocalin

0014-5793/02/$22.00
PII: S0014-5793(02)02606-6

as described previously [7]. Total RNA from *Arabidopsis* was isolated using the Tri-Reagent (Molecular Research Center) according to the manufacturer's protocol. Total RNA (7.5 μg) samples were mixed with ethidium bromide before electrophoresis on formaldehyde–agarose gels [8]. This allowed visual evaluation of RNA quantity and loads on gels. After electrophoresis, RNA was transferred to nitrocellulose membranes (Osmonics) in 20× saline sodium citrate (SSC). The filters were baked for 2 h at 80°C prior to hybridization with corresponding ^{32}P-labeled p*Tatil* and p*Attil* inserts. Filters were washed at 65°C with several buffer changes of decreasing SSC concentration (5–0.1×) and autoradiographed on Kodak MS films with MS intensifying screens (Kodak) at −80°C. Relative levels of *Tatil* and *Attil* mRNA transcripts were determined by densitometry scanning of the Northern blots using the ImageQuant 4.2 software (Molecular Dynamics).

Analysis and sequence comparisons were carried out with programs available on the ExPASy Molecular Biology Server.

3. Results and discussion

A novel plant lipocalin, *Tatil*, was identified from our collection of cold-induced EST using the BLAST X software [5]. The full-length clone was then isolated from a wheat cDNA library and named *Tatil* for *T. aestivum* L. temperature-induced lipocalin. The longest open reading frame of the cDNA is 570 bp and encodes a protein of 190 amino acids (aa) with an ATG codon at nucleotide 115 and a stop codon at nucleotide 686. The calculated molecular mass is 22 kDa and its theoretical p*I* is 5.5.

Search in the GenBank™ ESTs database revealed high homology (74% identity, 83% similarity) with a putative protein from *A. thaliana* that we have named *At*TIL for *A. thaliana* temperature-induced lipocalin. The complete sequencing of this *Arabidopsis* clone revealed that the cDNA encodes a 186 aa protein. Sequence analysis revealed that the N-terminal portion of both wheat and *Arabidopsis* proteins possesses the three lipocalin SCRs. The SCR 1 region is located within aa 15–31 (GLDVARYMGRWYEIASF) in *Ta*TIL and within aa 12–28 (GLNVERYMGRWYEIASF) in *At*TIL and possesses the two conserved amino acids G and W [1,9] (Fig. 1). The SCR 2 of *Ta*TIL is found at the C-terminal portion of the protein within aa 105–119 (YWVLYVDDDYQYALV) while in *At*TIL it is found within aa 101–115 (YWVLYIDPDYQHALI) (Fig. 1). Generally SCR 2 contains a TDY triplet at the positions underlined [1,9]. In *Ta*TIL and *At*TIL only the central D is present. SCR 3 is also found at the C-terminal portion of both proteins within aa 129–144 (ILCRKTHIEEEVNQL) in *Ta*TIL and within aa 125–140 in *At*TIL (ILSRTAQMEEETYKQL) (Fig. 1). The conserved R residue that characterizes this fingerprint is present in both sequences [1,9].

Further sequence analysis of *Ta*TIL indicated the presence of only one cysteine at aa 130 and a putative *N*-glycosylation site at aa 60. On the other hand, the entire primary sequence of *At*TIL does not contain any cysteine, although the putative *N*-glycosylation site is found at aa 56. Putative C-terminal cleavage sites are predicted by several target peptide prediction programs (DGPI, PSORT [10], and SignalP [11]) to be at aa 172 in *Ta*TIL and at aa 168 in *At*TIL. Considering this putative cleavage site, the calculated molecular mass of the mature protein of wheat and *Arabidopsis* is 20 kDa with a p*I* of 5.2.

The homology search revealed that *Ta*TIL (accession no. AY077702) and its ortholog from *Arabidopsis* (accession no. AY062789) share significant homology with three evolutionarily related lipocalins: the human ApoD precursor (accession no. P05090), the *Escherichia coli* outer membrane lipoprotein Blc precursor (accession no. P39281), and the American grasshopper Lazarillo precursor (accession no. P49291) (Fig. 1).

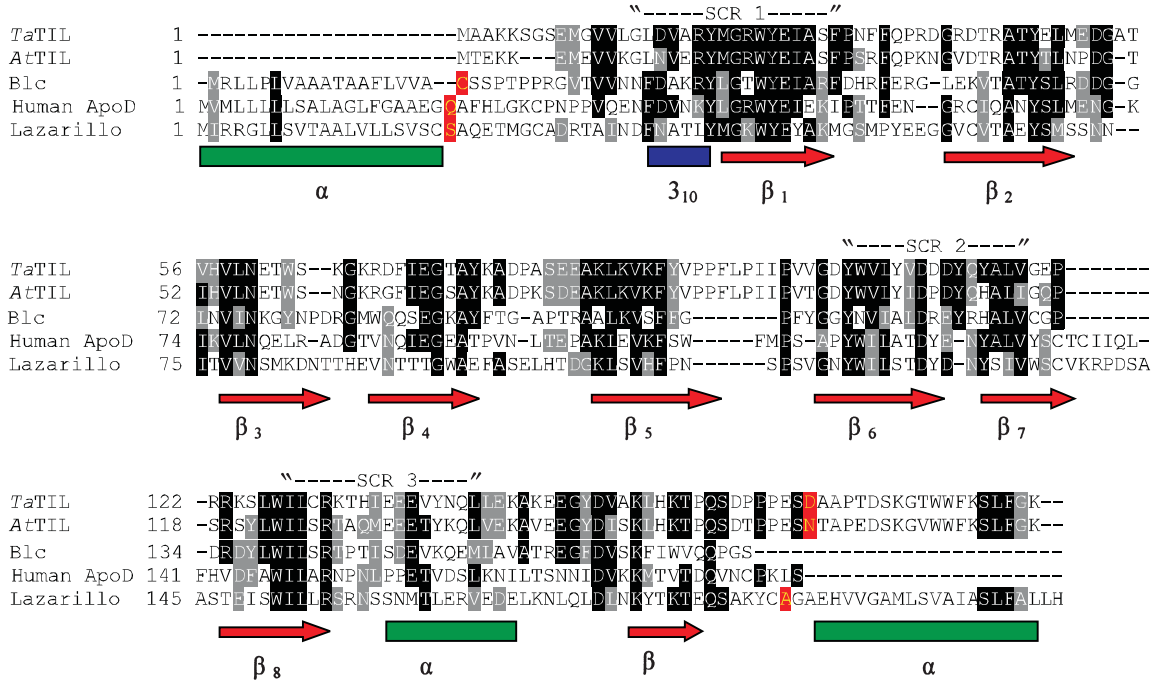

Fig. 1. Alignment of the deduced amino acid sequences of wheat *Ta*TIL and *Arabidopsis At*TIL with related lipocalins. Identical residues are in black and similar residues are in gray. The three SCRs that provide a signature for the lipocalins are indicated above. Putative cleavage sites are in red with yellow letters. The secondary structure is predicted from already published models [15,16] and analyzed with the Jpred2 program [23]. Red arrows, green rectangles and blue rectangle represent β-strands, α-helix, and 3_{10} α-helix respectively.

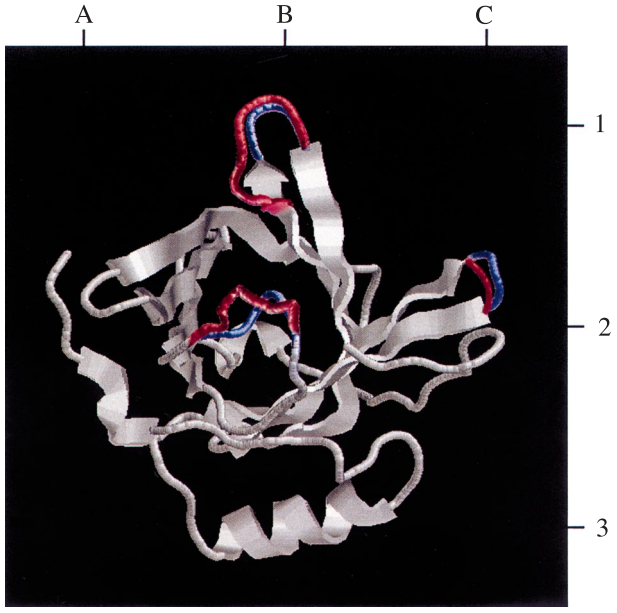

Fig. 2. Tertiary structure models of human ApoD and wheat *Ta*TIL. Tertiary structure analyses were carried out using the Swiss-Model program [24]. The lower blast limit was set at 0.00001 and the human ApoD model (PDB ID: 2APD) [15] was used as template. The initial result was then resubmitted through the optimizing mode of the program. The final result was then visualized using the Swiss-Pdb Viewer and the model was adapted according to sequence comparison. Differences between the wheat and the human models were superposed and colored. Sections of *Ta*TIL that differ from human ApoD are presented in red. Sections of human ApoD that differ from *Ta*TIL are in blue. Gray sections are common to both models.

They respectively share 29%, 31%, and 23% identity and 46%, 54% and 40% similarity with *Ta*TIL. Among all lipocalins, Blc, ApoD, and Lazarillo are the only ones known to be anchored to biological membranes [12]. Thus it is possible that *Ta*TIL and *At*TIL are also membrane-associated proteins.

The sequence analysis also revealed that *Ta*TIL and *At*TIL, like the *E. coli* Blc, are distinguished from most lipocalins by the absence of intramolecular disulfide bonds. However, they are potentially *N*-glycosylated like human ApoD and Lazarillo. When the three SCRs of these five proteins are aligned, the start methionines from *Ta*TIL and *At*TIL are positioned precisely at the cleavage sites of the N-terminal signal peptides of the three other proteins (Fig. 1). This alignment suggests that *Ta*TIL and *At*TIL do not possess an N-terminal signal peptide like *E. coli* Blc, human ApoD and Lazarillo. The N-terminal portion of *Ta*TIL is composed of hydrophilic residues followed by few hydrophobic residues. In *At*TIL, the hydrophobic section is even less accentuated. This profile does not fit the standard hydrophobic nature of the N-terminal signal peptide identified in ApoD, Blc and Lazarillo.

Like Lazarillo, *Ta*TIL and *At*TIL are longer than human ApoD and Blc at their C-terminal end and possess a similar putative cleavage site (Fig. 1). The hydrophobic C-terminal tail after the cleavage site enables Lazarillo to receive a glycosylphosphatidylinositol (GPI) anchor [13]. This may suggest that *Ta*TIL and *At*TIL also receive a GPI anchor. The GPI anchor is a post-translational addition of a lipid occurring in the endoplasmic reticulum lumen, which links proteins to the external face of the plasma membrane. This type of modifications has been reported in plants [14]. The fact that the *N*-glycosylation site is conserved between wheat and *Arabidopsis* TIL orthologs supports the possibility that those proteins are processed in the endoplasmic reticulum lumen.

Another type of attachment to the membrane can also be suggested for *Ta*TIL and *At*TIL. It has been proposed that human ApoD is associated with the external face of the membrane by a hydrophobic loop [12,15,16]. *Ta*TIL and *At*TIL also possess a hydrophobic stretch of seven amino acids that is inserted into a loop between two β-strands (Fig. 1). This insertion is not in the same loop as the human ApoD. It is between β-strands 5 and 6 instead of 7 and 8 (Fig. 2B1,C2).

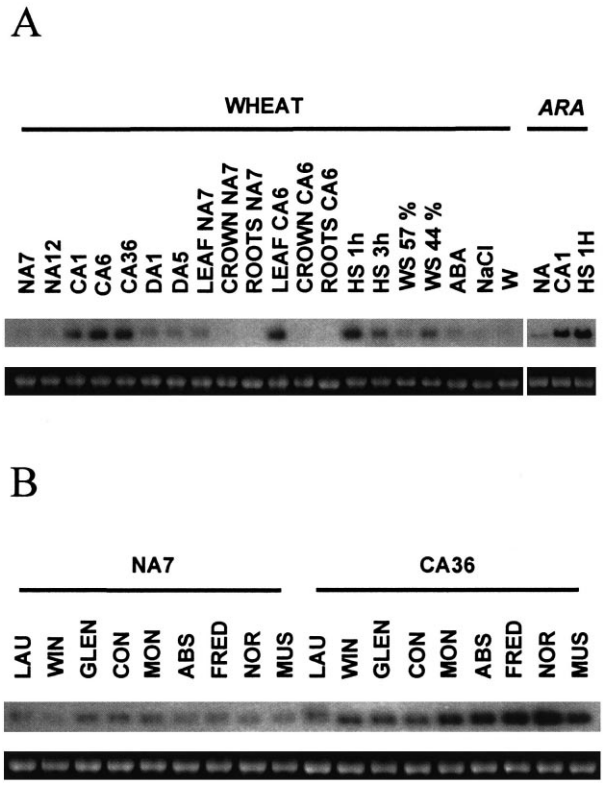

Fig. 3. Upregulation of *Tatil* and *Attil* during cold acclimation and heat-shock. A: Accumulation and tissue specificity of *Tatil* and *Attil* mRNAs under different stress conditions in winter wheat Fredrick and *Arabidopsis*. The 28S ribosomal band stained with ethidium bromide is included to show RNA loads (7.5 μg). NA7, NA12, non-acclimated plants grown for 7 and 12 days; CA1, CA6, CA36, cold-acclimated plants for 1, 6 and 36 days; DA1 and DA5, cold-acclimated plants (36 days) were deacclimated for 1 and 5 days; NaCl, plants treated with 300 mM NaCl for 18 h; ABA, plants treated with 0.1 mM ABA (Sigma) for 18 h; WS, water-stressed plants with a relative water content of 57% and 44%; HS, plants exposed to 40°C for 1 h and 3 h (heat-shock); W, wounding stress for 3 h. Leaf, crown and roots of non-acclimated plants (NA7) and 6 days cold-acclimated (CA6). *ARA*, *Arabidopsis* plants; NA, non-acclimated plants; CA1, cold-acclimated for 1 day; HS 1 h, exposed to 45°C for 1 h (heat-shock). B: Accumulation of *Tatil* mRNAs during cold acclimation in spring and winter wheat and other cereal species. Total RNA (7.5 μg) from shoots of: two spring wheat genotypes (*T. aestivum* L. cv Glenlea (Glen), LT_{50} (lethal temperature that kills 50% of the seedlings) −8°C; and cv Concorde (Con), LT_{50} −8°C), four winter wheat genotypes (*T. aestivum* L. cv Monopole (Mon), LT_{50} −15°C; cv Absolvent (Abs), LT_{50} −16°C; cv Fredrick (Fred), LT_{50} −16°C; and cv Norstar (Nor), LT_{50} −19°C), winter rye (*S. cereale* L. cv Musketeer (Mus), LT_{50} −21°C), oat (*A. sativa* L. cv Laurent (Lau), LT_{50} −6°C), barley (*H. vulgare* L. cv Winchester (Win), LT_{50} −7°C). NA, non-acclimated plants grown for 6 days; CA, cold-acclimated plants for 36 days.

However it is possible that this stretch favors the attachment to the plasma membrane. The model presented in Fig. 2 reveals a difference in the loop scaffold covering the large cup-shaped cavity that characterizes the lipocalins (Fig. 2B2). The loop scaffold in *Ta*TIL and *At*TIL is two amino acids longer than human ApoD and has a proline moiety at positions 32 and 29 respectively (Fig. 1). Those modifications may suggest that the plant TIL has a different binding specificity.

Northern blot analysis revealed that the *Tatil* transcripts accumulate to high levels upon exposure to low temperature (10-fold) and heat-shock treatments (10-fold) and to a lesser extent by water stress (3.5-fold). Abscisic acid (ABA), high salt and wounding treatments have no measurable effect (Fig. 3A). The *Tatil* transcripts accumulate gradually to a maximum level after 36 days of cold acclimation. Upon deacclimation, the level of transcripts returned to those seen in the control non-acclimated plants. The accumulation of *Tatil* transcripts in wheat was found to be tissue-specific, as they were detected only in cold-acclimated leaves (Fig. 3A). Transcripts accumulation of *Attil* revealed that the dicot ortholog is also induced by low temperature (six-fold) and heat-shock treatments (nine-fold) (Fig. 3A). *Arabidopsis* genome sequence analysis (accession no. AB024029) revealed that the promoter of *Attil* does not contain any known heat-shock responsive element although two low temperature responsive elements are found [17,18].

RNA blot hybridization studies also demonstrated that cold acclimation induced the accumulation of the *Tatil* transcripts in both less tolerant and hardy wheat (Fig. 3B). However, this increase was greater in the hardy winter cultivars. Low levels of expression are also found in oat and barley, two less cold tolerant species (Fig. 3B). This difference of accumulation indicates that the *Tatil* expression is correlated with the plant's capacity to develop freezing tolerance.

Temperature stresses are known to induce membrane injuries. The membrane-anchored lipocalins (Blc, ApoD, Lazarillo and possibly *Ta*TIL and *At*TIL) all appear to be expressed in response to conditions that cause membrane stresses [12], which suggest a biological role in membrane biogenesis and repair under severe stress conditions.

*Ta*TIL and *At*TIL like human ApoD may possess a wide variety of potential ligands of different structures and functions. Mammalian ApoD is reported to bind arachidonic acid, bilirubin, steroid hormones (progesterone and pregnenolone) and cholesterol [19]. It is interesting to mention that plants also synthesize a wide variety of steroid hormones called brassinosteroids. Treatment with 24-epibrassinolide, a brassinosteroid, increases the tolerance to heat and cold stresses in plants [20]. The enhanced resistance to temperature stress was attributed to membrane stability and osmoregulation. These results suggest that part of the temperature responses in plants may involve brassinosteroids as signaling molecules to elicit the expression of steroid binding proteins such as lipocalin. It is also known that sterol insertion in the plasma membrane increases its fluidity at low temperature and maintains the phospholipids order at high temperature [21]. *Ta*TIL may be involved in the transport of those sterol molecules to the membrane in response to stress conditions.

The specific expression of *Ta*TIL in leaf suggests that the protein may play a role in the chloroplast function during temperature stresses. However, there is no evidence of the presence of a chloroplast signal peptide to confirm this assumption. Recent work on the OEP7 protein from spinach demonstrated that the interaction of this protein with the outer membrane of the chloroplast is independent of a classical cleavable targeting signal or membrane-channel protein [22]. It is possible that the *Ta*TIL is anchored to the outer chloroplast membrane in a similar manner under temperature stresses.

References

[1] Flower, D.R. (1996) Biochem. J. 318, 1–14.
[2] Bugos, R.C., Hieber, A.D. and Yamamoto, H.Y. (1998) J. Biol. Chem. 273, 15321–15324.
[3] Danyluk, J., Carpentier, E. and Sarhan, F. (1996) FEBS Lett. 389, 324–327.
[4] Houde, M., Danyluk, J., Laliberte, J.F., Rassart, E., Dhindsa, R.S. and Sarhan, F. (1992) Plant Physiol. 99, 1381–1387.
[5] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[6] Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S. and Thomashow, M. et al. (1994) Plant Physiol. 106, 1241–1255.
[7] Danyluk, J. and Sarhan, F. (1990) Plant Cell Physiol. 31, 609–619.
[8] Rosen, K.M. and Villa-Komaroff, L. (1990) Focus 12, 23–24.
[9] Flower, D.R., North, A.C. and Attwood, T.K. (1993) Protein Sci. 2, 753–761.
[10] Nakai, K. and Horton, P. (1999) Trends Biochem. Sci. 24, 34–36.
[11] Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Int. J. Neural. Syst. 8, 581–599.
[12] Bishop, R.E. (2000) Biochim. Biophys. Acta 1482, 73–83.
[13] Ganfornina, M.D., Sanchez, D. and Bastiani, M.J. (1995) Development 121, 123–134.
[14] Morita, N., Nakazato, H., Okuyama, H., Kim, Y. and Thompson Jr., G.A. (1996) Biochim. Biophys. Acta 1290, 53–62.
[15] Peitsch, M.C. and Boguski, M.S. (1990) New Biol. 2, 197–206.
[16] Bishop, R.E., Penfold, S.S., Frost, L.S., Holtje, J.V. and Weiner, J.H. (1995) J. Biol. Chem. 270, 23097–23103.
[17] Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Nucleic Acids Res. 27, 297–300.
[18] Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. and Tabata, S.J. (2000) DNA Res. 7, 31–63.
[19] Rassart, E., Bedirian, A., Do Carmo, S., Guinard, O., Sirois, J., Terrisse, L. and Milne, R. (2000) Biochim. Biophys. Acta 1482, 185–198.
[20] Clouse, S.D. and Sasse, J.M. (1998) Annu. Rev. Plant Physiol. Plant Mol. 49, 427–451.
[21] Demel, R.A. and De Kruijff, B. (1976) Biochim. Biophys. Acta 457, 109–132.
[22] Schleiff, E., Tien, R., Salomon, M. and Soll, J. (2001) Mol. Biol. Cell 12, 4090–4102.
[23] Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J. (1998) Bioinformatics 14, 892–893.
[24] Peitsch, M.C. (1996) Biochem. Soc. Trans. 24, 274–279.