This is to certify that the

dissertation entitled

STUDIES ON THE REGULATION AND PROTEIN PRODUCTS OF THE COR15 GENE FAMILY IN ARABIDOPSIS THALIANA

presented by

Kathy Suzanne Wilhelm

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Genetics

Major professor

Date July 29, 1996

DATE DUE	DATE DUE	DATE DUE
OCT 19 1996 282		
06176 05 NOV 26 2003		

STUDIES ON THE REGULATION AND PROTEIN PRODUCTS OF THE *COR*15 GENE FAMILY IN *ARABIDOPSIS THALIANA*

By

Kathy Suzanne Wilhelm

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Program in Genetics

1996

ABSTRACT

STUDIES ON THE REGULATION AND PROTEIN PRODUCTS OF THE *COR*15 GENE FAMILY IN *ARABIDOPSIS THALIANA*

By

Kathy Wilhelm

During cold acclimation in *Arabidopsis thaliana*, several families of genes are induced, but the mechanism(s) by which they are induced remains unknown. A genomic clone containing the *COR*15 gene family was isolated for the purpose of examining the cold-regulation of its promoter. This clone was found to contain two tandem members of the family, *COR*15a and *COR*15b. Their predicted coding regions are 82% identical and both are transcriptionally regulated by low temperature. A structural model of the protein encoded by *COR*15b is proposed. The promoter of *COR*15a fused to GUS was then used as a screenable marker in an attempt to find signal transduction mutants aberrant in the regulation of *COR*15a. The only mutants found showed aberrant regulation of the transgene, but normal induction of the endogenous gene, indicating that the mutations were not in the desired pathway. Ways to improve the method are discussed.

An examination of the temperature induction profiles of four *COR* gene families showed that all are induced between 19 and 16°C, and that message levels gradually increase as the temperature is lowered and level off at about 8°C. For *COR*15a, this profile appears to be promoter-based, since message levels of a *COR*15a/GUS promoter fusion gave the same profile as the endogenous gene. The temperature induction profile was the same whether the temperature was progressively lowered two or three degrees at a time or was shifted by as much as 18°C at once, suggesting that the signal involves the temperature *per se* and not the change in temperature. Also, mutants deficient in polyunsaturated fatty acids in the chloroplast or plasma membrane show the same induction profile as wildtype plants. This fails to support the idea that an alteration in membrane fluidity may be involved in the induction of *COR* gene families.

DEDICATION

To the One around Whom my life revolves,
without Whom not a word could have been written,
the King of Kings and Lord of Lords.
May all honor and praise be His in this and in everything I do.

ACKNOWLEDGEMENTS

Thanks are due to my professor, Dr. Mike Thomashow, for providing a stimulating, amicable research environment and for offering insight and guidance throughout this project. I also appreciate the guidance and challenge I have received from my guidance committee, Dr. Pam Green, Dr. Ken Poff and Dr. Steve Triezenberg. I offer many thanks to Dr. Leslie Kuhn for her infectious enthusiasm and for the expertise in protein modelling that made the model of COR15bm possible, and to Pappan (Kailla Padmanabhan) for his gracious assistance in the operation of unix-based computers. I am also grateful to Michelle Marshall and Bob Gifford who were a tremendous help in screening thousands of mutagenized plants. Thanks are due to Dave Horvath, for showing me the ropes when I entered the laboratory, to Stokes Baker for working with me in the analysis of *COR*15a and for providing the transgenic plants I used in the mutagenesis experiment, to Leonard Bloksberg for his assistance in analyzing the plants transgenic for *COR*15b promoter constructs and to Sarah Gilmour for her editorial skills and general knowhow around the lab. I am grateful to my colleagues Eric Stockinger, Kevin O'Connell, Chentao Lin, Todd Cotter, Dan Zarka, Nancy Artus, Weiwen Guo, Brett McLarney, Deane Lehman, Ann Gustafson, Beth Seymour, Kirsten Jaglo-Ottosen, to the many students (graduate

and undergraduate) who have worked in Dr. Thomashow's laboratory, and to the members of Dr. Rebecca Grumet's laboratory for their discussions, assistance and general friendliness. Finally, I wish to thank my Godbrother, Ralph Benedetto, for his support, encouragement and editorial assistance.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1:

Literature Review

Introduction

In late summer or early autumn in northern climes, the cry goes forth "There's going to be a frost tonight! Cover your tomatoes!" No one worries about protecting their carrots or columbine. These plants and many others survive not only the light frost, but even the subzero weather of deep winter without intervention. What is the difference? Why do snowdrops and crocuses come up through the snow and bloom successfully, while cherry blossoms are sterilized by an ill-timed frost?

While the emphasis of this work is on an examination of gene regulation in *Arabidopsis thaliana* in response to low temperature, it is worthwhile to place this within the larger context of what happens in plants exposed to low temperature. The questions presented can only be partially answered. A brief examination of the literature quickly reveals their complexity. The response of a given plant to low temperature depends on the organ, the stage of growth, the amount of light,

the humidity, the temperature, the time spent at low temperature (Raison and Orr, 1990), the rate of cooling (Minorsky, 1989) the plant's climate of origin and whether the plant has been hardened.

With respect to low temperature, plants fall into two major classes that correspond to their climate of origin. Those from tropical or subtropical climates, like tomatoes and cucurbits, are chilling-sensitive. There is substantial variability in the susceptibility of chilling-sensitive plants, but when temperatures drop below a threshold level (about 10°C for most chilling-sensitive species), they suffer damage ranging from loss of vigor to death. Of the many physiological changes that occur at low temperature, the main causes of injury in chilling-sensitive plants appear to be dehydration from a drop in root pressure (Minorsky, 1985) and the failure of stomata to close (Patterson and Reid, 1990), the production of toxic oxygen species due to impaired photosynthesis and respiration (Purvis and Shewfelt, 1993), and the increased permeability of cellular membranes (Murata, 1990). Which low temperature response(s) is primary and which is secondary is not yet completely clear, although Lyons (1973) and Raison (1973) propose that a low temperature-induced change in the physical state of membranes in chilling-sensitive plants leads to the other symptoms observed. Minorsky (1985) suggests that increases in cytosolic calcium levels might be responsible for the symptoms of chilling injury. Subsequent studies with transgenic *Nicotiana plumbaginifolia*, a chilling-

sensitive plant, have shown that cold shock to 0 or 5°C does cause a transient increase in the intracellular calcium concentration (Knight et al., 1991). However, a shock to 10°C, a temperature at which many chilling sensitive plants are susceptible to growth inhibition or damage, did not cause an increase in intracellular calcium in these experiments (Knight et. al., 1991).

It is also known that many chilling-sensitive plants can become more resistant to chilling temperatures if they are chill- or drought-hardened, and that most show delayed injury when chilled at 100% relative humidity (Wilson, 1979).

Plants of temperate or arctic origin are chilling-tolerant, although certain organs or stages of life may be chilling-sensitive (Bramlage and Meir, 1990). Being poikilothermic organisms, they have little opportunity to avoid low temperature, so resistant plants must either tolerate or be insensitive to its numerous effects. With respect to the water stress associated with chilling injury, Markhart (1986) has shown that when detopped roots of bean (sensitive) and spinach (resistant) were chilled, the bean lost 90% of its root conductance while the spinach lost 80%. Within eight hours, the bean had only recovered to 30% of its original capacity, while the spinach was up to 70%. Thus, the most severe water stress in spinach was transient. Long-term water stress was avoided. Resistant plants may also avoid the damage caused by toxic oxygen species. Purvis and Shewfelt (1993) report that "cold-resistant cultivars and chilling-resistant tissues generally develop a greater potential for respiratory electron flux through the alternative

pathway than do cold sensitive cultivars and tissues." Electrons directed through the alternative pathway are not available to create toxic oxygen species. Finally, differences in the composition of cellular membranes may allow resistant plants to avoid membrane permeability. A high proportion of saturated phosphatidylglycerol (PG) and sulfoquinovosyldiacylglycerol (SQDG), glycerolipids with phase transition temperatures above 30°C, is well correlated with chilling-sensitivity (Murata and Nishida, 1990).

Work with *Arabidopsis* mutants deficient in membrane fatty acid lipid unsaturation has also suggested that membrane composition affects chilling tolerance. Wildtype *Arabidopsis* is not injured by low temperature, although its growth rate declines considerably and it has a higher content of chlorophyll when grown at 5°C (Hugly and Somerville, 1992). While several *fad* mutants are morphologically indistinguishable from wildtype plants at either normal growth temperatures or under chilling stress, *fad*B (*fad*5) and *fad*C (*fad*6) grown at 5°C have a reduced growth rate and chlorophyll content compared to wildtype *Arabidopsis* grown at 5°C. Leaves that develop at 5°C in these two mutants are chlorotic (Hugly and Somerville, 1992). A third *fad* mutant, *fad*2-2, develops necrotic lesions and eventually dies if maintained at 6°C for longer than ten days (Miquel et al., 1993).

While the molecular mechanisms behind the differences between chilling sensitive and chilling-resistant plants are still under investigation, Murata and colleagues have shown that a glycerol-3-phosphate acyltransferase isolated from pea and spinach, which are chilling resistant, preferentially acylated the unsaturated fatty acid 18:1 to the *sn*-1 position of glycerol 3-phosphate. The same enzyme isolated from chilling-sensitive squash hardly discriminated between unsaturated 18:1 and saturated 16:0 (Murata and Nishida, 1990). They then transformed tobacco with the glycerol-3-phosphate acyltransferase gene isolated from squash, a more chilling-sensitive plant than tobacco, or *Arabidopsis*, which is chilling tolerant, and found that the tobacco plants transformed with the squash enzyme had lower levels of cis-unsaturated fatty acids in PG and were more chilling-sensitive than wildtype tobacco. The opposite was true of tobacco transformed with the *Arabidopsis* gene (Murata et al., 1992). Thus it has been shown directly and in vivo that the selectivity of a single enzyme can affect membrane composition and alter chilling tolerance. It remains to be learned why and how this is so.

It has also been noted that tobacco plants overexpressing chloroplastic Cu/Zn superoxide dismutase (SOD) from pea are less sensitive to chilling injury than wildtype tobacco (Gupta et al., 1993). Whether SOD from chilling-resistant pea is superior to tobacco isoforms under low temperature conditions or whether the

simple overexpression of SOD caused this lessening of susceptibility is not known.

Some chilling-resistant species are also freezing-tolerant, many of them able to increase in freezing tolerance in response to low nonfreezing temperatures. Plants damaged due to freezing typically appear flaccid and water-soaked, suggesting that freezing injury is mainly due to the compromise of cell membrane integrity. Depending on the nature of the freeze and on the properties of the cells involved, this may be manifested as intracellular ice formation; as expansion-induced lysis, which is caused by the inability of the thawing cell to properly rehydrate; or as the loss of osmotic responsiveness (Steponkus, 1984).

Kendall et al. (1989) suggest that membrane disruption may be due to free radicals produced during freezing, since they could detect free radical production during the freezing of winter wheat, and acclimated plants were more resistant to applied free radicals as well as to freezing than nonacclimated plants. Also, the constitutive expression of an extra superoxide dismutase in alfalfa resulted in one transgenic line that had both greater resistance to a free radical-producing herbicide and enhanced freezing tolerance (McKersie et al., 1993).

Intracellular ice formation, whether it be due to intracellular nucleation or to penetration of the cell by an external ice crystal, is generally believed to be instantly lethal unless the rates of freezing and thawing are rapid enough that the ice crystals are very fine (Sakai and Sugawara, 1978). The intact plasma membrane is thought to be an effective barrier to seeding by extracellular ice (Chambers and Hale, 1932), and the work of Steponkus and Dowgert (1983) suggests that intracellular seeding may be a result of an alteration of the plasma membrane, since mechanical failure of the plasmalemma can be observed prior to intracellular ice formation.

More commonly, ice formation is extracellular. The solute concentration of the apoplast is lower than that of the cytoplasm, with the result being that the cytoplasm has a greater freezing point depression (Guy, 1990). Once the extracellular water begins to freeze, however, the water potential outside the cell decreases. Ice has a lower water potential than water at the same temperature and solutes excluded from the growing ice crystal lower the apoplastic water potential further, so water diffuses out of the cell (Thomashow, 1994). Thus the cell suffers from dehydration as well as from the temperature itself. Plants frozen to -10°C lose more than 90% of their osmotically active water (Steponkus and Lunch, 1989). When the cells are rehydrated, water moves back into the cell, and if the cells are unable to accommodate the influx, "expansion-induced lysis" is the result (Steponkus, 1984).

Loss of osmotic responsiveness occurs after nonacclimated protoplasts are frozen below -5°C, but can be mimicked by osmotic dehydration in the absence of ice formation. This loss is associated with changes in the ultrastructure of the plasma membrane, including lamellar to hexagonal-II phase transitions (Steponkus and Lynch, 1989), in which lipids are reoriented. Instead of forming a bilayer, the lipids are arranged in long cylinders with their polar head groups in an aqueous core (Steponkus, 1984).

An increase in freezing tolerance, therefore, acquired in a process commonly referred to as "cold acclimation," must prepare the cell for avoidance or tolerance of dehydration and rapid rehydration, the formation of intracellular ice, and the loss of osmotic responsiveness. This being the case, it makes sense that drought-hardening, which can cause an increase in chilling tolerance of chilling-sensitive plants, can also bring about an increase in freezing tolerance in chilling-resistant cabbage (Cox and Levitt, 1976), spinach (Guy et al., 1992), wheat and rye (Siminovitch and Cloutier, 1983).

It has been seen that protoplasts isolated from cold acclimated plants experience intracellular ice formation at lower temperatures than those from nonacclimated plants (Dowgert and Steponkus, 1983). It has also been observed that protoplasts isolated from cold acclimated rye, unlike those from nonacclimated rye, were not susceptible to expansion-induced lysis after having been frozen to

-5°C (Steponkus, 1984). Finally, protoplasts isolated from acclimated rye become osmotically unresponsive at much lower temperatures than those from nonacclimated rye, and even then, they undergo different membrane phase transitions than do their nonacclimated counterparts (Fujikawa and Steponkus, 1990). This difference in membrane phase transitions during freezing has been confirmed in studies using leaf sections taken from acclimated or nonacclimated rye (Webb and Steponkus, 1993). Thus, the avoidance of intracellular ice formation and of loss of osmotic responsiveness and the tolerance of rapid rehydration appear to be operative in frozen acclimated plants.

The physical properties of the tolerance of the dehydration and rehydration in acclimated plants have been clarified by Steponkus and colleagues (Steponkus, 1984). They demonstrated that protoplasts isolated from nonacclimated rye form endocytotic vesicles which bud off from the plasma membrane as the cells shrink during freeze-induced dehydration. Upon thawing, rehydration results in intolerable osmotic pressure because the vesicles are not reincorporated into the membrane and the protoplasts burst. Protoplasts from acclimated rye, however, form exocytotic extrusions as they dehydrate. These extrusions remain in association with the membrane so that when the cells regain water, the extrusions are reincorporated into the membrane and the cells do not lyse. It has also been shown that a change in lipid composition is sufficient to favor either endocytotic vesiculation or exocytotic extrusions (Steponkus and Lynch, 1989).

Indeed, the membrane composition of *Arabidopsis* is such that even protoplasts isolated from nonacclimated plants are resistant to expansion-induced lysis (Umera et al., 1995). Whether whole plant cells behave in this manner remains to be seen.

The process of cold acclimation, by which these changes are brought about, involves many physiological alterations. The lipid composition of cellular membranes changes (Lynch and Steponkus, 1987, Umera et al., 1995), resulting in the differences in membrane response to freezing already described. There are increased levels of sugars, soluble proteins, proline and other organic acids (Sakai and Larcher, 1987). These may be involved in ameliorating the effects of freeze-induced dehydration and/or in reducing its severity. New isozymes are formed (Guy, 1990) and changes in the expression of a modest number of genes are seen (Thomashow, 1993).

Cold-inducible genes

While the overall pattern of gene expression in plants exposed to low temperature does not drastically change (Gilmour et al., 1988), still, quite a number of *COR* (*co*ld *r*egulated) genes have been identified from numerous plants. For the sake of brevity, and because the plant examined in this work is resistant to chilling damage, only genes induced in chilling-tolerant plants will

be considered here. The functions of some low temperature-inducible genes have been demonstrated experimentally or inferred based on sequence comparisons with genes of known identity.

The usefulness of some of the genes found to be induced at low temperatures is readily explainable. Increased levels of sucrose synthetase in wheat (Newsted et al., 1991, Marana et al., 1990 and Crespi et al., 1991) and of phosphoenolpyruvte carboxykinase in rapeseed (Saez-Vasquez et al., 1995) may be involved in the increase in soluble sugar content thought to ameliorate the dehydration stress imposed on frozen cells. The accumulation of extensin mRNA in acclimated pea seedlings is thought to lead to a buildup of the extensin protein, which may protect cell walls from collapse under the extreme dehydration pressures caused by freezing (Weiser et al., 1990). This protein contains a large portion of the hydroxyproline in the cell, and levels of hydroxyproline in cell walls increased during the same period that mRNA levels of extensin increased, suggesting that extensin was being incorporated into cell walls and enhancing their rigidity (Weiser et al., 1990). Antifreeze proteins found at increased levels in acclimated winter rye could help prevent recrystallization of extracellular ice, during which crystals large enough to cause physical damage to tissues and cells could form (Hon et al., 1995). Lipid transfer proteins induced in acclimated barley could play a role in the alteration of lipid content seen in cell membranes at low temperature (Hughes et al., 1992, White et al., 1994).

With respect to general metabolism, heat shock proteins and heat shock cognate proteins found in acclimated Brassica napus (Krishna et al., 1995) or spinach (Neven et al., 1992, Li et al., 1994, Anderson et al., 1994) may be involved in protecting proteins from low temperature denaturation or in helping proteins to refold. Alcohol dehydrogenase induced at low temperature is thought to play a role in the shift to anaerobic metabolism that occurs when respiration in inhibited (Jarillo et al., 1993). Phenylalanine ammonia-lyase and chalcone synthase, enzymes involved in the production of anthocyanins, are also induced by exposure to low temperature (Levya et al., 1995). Anthocyanins are believed to act as light-screening pigments (Levya et al., 1995) and may help decrease the amount of toxic oxygen species produced by cold-impaired photosynthesis.

The induction of genes that could be involved in the regulation of translation or in signal transduction pathways active at low temperature was also observed. In barley, a translation elongation factor 1α is inducible in cold-treated plants (Dunn et al., 1993). In addition, several protein kinases, including a novel protein kinase in wheat (Holappa and Walker-Simmons, 1995), two calcium-dependent protein kinases in alfalfa (Monroy and Dhindsa, 1995), a mitogen-activated protein kinase kinase kinase in *Arabidopsis* (Mizoguchi et al., 1996) and two genes with high sequence similarity to ribosomal-protein S6 kinases, also in *Arabidopsis* (Mizoguchi et al., 1995), are inducible by low temperature. Finally, two genes showing high identity with 14-3-3 proteins show low temperature

inducibility in *Arabidopsis* (Jarillo et al., 1994). Thought to be involved in the regulation of protein kinases, 14-3-3 proteins are likely to be involved in signal transduction. One member of the family has been found to be part of the protein complex that binds to the g-box promoter element in *Arabidopsis* (Lu et al., 1992). This supports the idea that 14-3-3 proteins may have a role in gene regulation in plants.

The purpose of a few other identified cold-responsive genes in low temperature acclimation or metabolism is less obvious. In *Arabidopsis*, message levels of one member of the β-tubulin gene family increase while levels of four other members of the family decrease and still others remain the same (Chu et al., 1993). Low temperature does affect the integrity of microtubules (Pihakaski-Maunsbach and Puhakainen, 1995), but their role in cold tolerance is still unknown. Also in *Arabidopsis*, Williams et al. (1994) have found a potential thiol protease is induced in the cold. They suggest that it may be involved in the degradation of proteins denatured by low temperature, in the degradation of storage proteins to alter the osmotic potential of the cell or in the stress-induced proteolytic activation of enzymes. Another low temperature-inducible gene which may assist in osmotic adjustment is the 70 kd subunit of tonoplast ATPase, which has been isolated from winter *Brassica napus* (Orr et al., 1995).

In addition to these genes of known or predicted identity, there are many with little or no sequence identity to genes of known function. These include HVA1 (Sutton et al., 1992), *blt*101, *blt*1015, *blt*63, *blt*49, *blt*410, *blt*14, *blt*411, *blt*801 (Dunn et al., 1994), COR14 (Crosatti et al., 1995), Dhn1 and Dhn2 (van Zee et al., 1995) from barley, *cor39/Wcs120* from wheat (Guo et al., 1992, Houde et al., 1992), and possibly pBGA12, pBGA56, pBGA85 and pBGA25 from bromegrass (Lee and Chen, 1993). In leguminous plants, ELIP from green pea (Adamska and Kloppstech, 1994), GAB-8 and GAB-9 from chick pea (Colorado et al., 1994), *cas15* (Monroy et al., 1993a), *cas18* (Wolfraim et al., 1993) and *msa* CIC (Castonguay et al., 1994) from alfalfa have been found to be responsive to low temperature. In spinach, the gene or genes encoding CAP85 show cold-inducibility (Neven et al., 1993). Low temperature-responsive genes in crucifers include BN19, BN26, BN115 (Weretilnyk et al., 1993), BnC24A, BnC24B (Saez-Vasquez et al., 1993), BN28 (Boothe et al., 1995) and *btg*-26 (Stroehr et al., 1995) from *Brassica napus*, and *rab18* (Lang et al., 1992), *lti30* (Welin et al., 1994), *kin1* (Kurkela and Franck, 1990), *COR*6.6/*kin2* (Gilmour et al., 1992, Kurkela and Borg-Franck, 1992), *COR*15a (Lin and Thomashow, 1992), *COR*15b (Wilhelm and Thomashow, 1993), *COR*47 (Gilmour et al., 1992), *lti45/lti29* (Welin et al., 1994, Welin et al., 1995), *COR*78/*rd29A/lti78* (Horvath et al., 1993, Yamaguchi-Shinozaki and Shinozaki, 1993, Nordin et al., 1993) and *rd29B/lti65* (Yamaguchi-Shinozaki and Shinozaki, 1993, Nordin et al., 1993) from *Arabidopsis*. Many of the genes listed show sequence similarity to dehydrins or late embryogenesis

abundant genes, which are prevalent in plants during conditions of water stress, but the functions of these genes are still a matter of speculation.

Work done with the *COR* gene families of *Arabidopsis* has shown that the known members of families *COR*6.6 (*kin1* and *COR*6.6), *COR*15 (*COR*15a and *COR*15b), *COR*47 (*COR*47 and *lti45/lti29*) and *COR*78 (*COR*78/*rd29A*/*lti78* and *rd29B*/*lti65*) are all inducible by low temperature (Thomashow, 1994). The mRNA levels of the gene families remain high as long as plants are kept at low temperature, declining to control levels within eight hours of deacclimation (Hajela et al., 1990). At least one member of each family is induced by drought stress and by the exogenous application of abscisic acid (ABA), but none respond to heat shock. All are very hydrophilic and remain soluble upon boiling.

Of these *Arabidopsis COR* genes, the most is known about *COR*15a. Promoter deletion studies have shown that its expression is regulated at the promoter level by low temperature, drought and ABA and suggest that the promoter elements responsible for this regulation are located between -305 and -78, relative to the start of transcription (Baker et al., 1994). Indeed, an element found in this region, CCGAC, contains the five core bases of the DRE, an element from the *COR78*/*rd29A*/*lti78* promoter identified as being sufficient for cold and drought inducibility (Yamaguchi-Shinozaki and Shinozaki, 1994). Expression of *COR*15a

at low temperature is found in most tissues of the plant, except for the roots and ovaries. There is also constitutive expression of a *COR*15a/GUS promoter fusion construct in anthers (Baker et al., 1994).

The protein encoded by *COR*15a, COR15a, is targeted to the chloroplast. This might be expected from its expression pattern, since it is mainly found in green tissues. In the process of import into the chloroplast, COR15a is cleaved from its original molecular weight of 14.7 kd to the 9.4 kd mature form, COR15am (Lin and Thomashow, 1992). The mature form of the protein is acidic, with a pI of 4.6 (Gilmour et al., 1996). The function of COR15am within the chloroplast is not yet known, but it has recently been shown that transgenic plants carrying the *COR15a* coding region driven by a constitutive promoter, which express *COR*15a at normal growth temperatures, show increased freezing tolerance. The efficiency of photosystem II was less sensitive to freezing, as measured by the Fv/Fm ratio, than in non-transgenic controls or in transgenic plants containing an unrelated coding region, and protoplast survival was enhanced (Artus et al., in preparation). The mechanism for this increased freezing tolerance has yet to be elucidated.

Low temperature signal transduction

Among the many questions yet to be answered is that regarding the mechanism by which cold-regulated genes are induced. Nor is it clear that any or all cold-regulated genes are induced by the same signal. Mantyla et al. (1995) have shown that in mutants deficient in or insensitive to ABA, RAB18 (the protein product of *rab18*) is no longer cold-inducible, while LTI78 (the protein encoded by *lti78/COR78*) is responsive to low temperature, suggesting that the low temperature induction of RAB18 requires the presence of ABA, while that of LTI78 does not.

Given the many changes that occur within the cell upon exposure to low temperature, there are many possible signals. A signal cascade could be initiated by a change in water potential, by an increase in the amount of reactive oxygen species within the cell, by the conformational change of a cold-sensitive protein, by an alteration in the properties of cellular membranes or by some other cellular consequence of low temperature. The signal could be as simple as a single protein that is both sensor and transcriptional activator, as is true of sterol regulatory element-binding protein 1 (*SREBP-1*) in mammals (Wang et al., 1994) or as complex as a multi-component, branched cascade. Membrane fluidity appears to be instrumental in the induction of the low temperature-responsive gene *desA* from the cyanobacterium *Synechocystis* PCC6803. This gene

is responsive to a change in temperature of greater than five degrees Celsius (Los et al., 1993), but is also responsive to palladium-catalyzed hydrogenation of the plasma membrane (Vigh et al., 1993). Whether membrane fluidity or any membrane-localized low temperature response alters the expression of *COR* genes in plants is still unknown.

Other factors that may play a role in signal transduction include the levels of ABA, the intracellular calcium concentration and the phosphorylation state of proteins. ABA has been implicated in the process of acclimation because ABA levels rise in response to low temperature in potato (Chen et al., 1983), winter wheat (Lalk and Dorffling, 1985) and spinach (Guy and Haskell, 1988) and because exogenous application of ABA can cause increased freezing tolerance in potato (Chen et al., 1979), alfalfa (Mohapatra et al. 1989), *Arabidopsis* (Lang et al., 1988), and in cell suspension cultures of winter wheat, winter rye and bromegrass (Chen and Gusta, 1983), while some of the ABA mutants of *Arabidopsis* are impaired in freezing tolerance (Heino et al., 1990, Gilmour and Thomashow, 1991). As was mentioned earlier, ABA appears to be needed for the cold-regulation of RAB18, but not LTI78 (Mantyla et al., 1995). Thus, it seems to be required for some, but not all, of the changes that occur during low temperature.

Calcium has been implicated in low temperature metabolism in a number of ways. In onion, tension-dependent activity of calcium-selective cation co-channels increased as the temperature was lowered from 25°C to 6°C (Ding and Pickard, 1993). A transient increase in cytosolic calcium levels was seen in tobacco (*Nicotiana plumbaginifolia*) (Knight et al., 1991) and in *Arabidopsis* (Knight et al., 1996) in response to cold shock. In tobacco, the cold shock temperature had to be 5°C or lower for the increase in calcium levels to be detectable (Knight et al., 1991), while *Arabidopsis* was only tested using ice water for the cold shock (Knight et al., 1996). A calcium influx was also seen in alfalfa below 15°C (Monroy and Dhindsa, 1995). That calcium levels may actually be involved in the alteration of gene expression is suggested by work done in chick pea and alfalfa (Colorado et al., 1994, Monroy and Dhindsa, 1995). In chick pea, genes that are responsive to low temperature, heat shock, ABA and the osmotic stress applied by NaCl or polyethylene glycol (PEG) were also up-regulated in the presence of 0.5 mM $CaCl_2$ (Colorado et al., 1994). More extensive experiments performed with alfalfa cell suspension cultures examined the inducibility of *cas15* and *cas18*, which are inducible by low temperature, but not by ABA, heat shock, water stress (imposed by a solution of PEG-6000) or wounding (Mohapatra et al., 1989). Inhibitors which blocked the influx of external calcium also inhibited the induction of the *cas* genes at low temperature, while the addition of a calcium ionophore or a calcium channel agonist resulted in *cas* gene induction at 25°C (Monroy and Dhindsa, 1995).

Cascades of phosphorylation and dephosphorylation reactions are a common theme in signal transduction in animals and are becoming common in the plant literature as pathways are being unraveled (Raghuram and Sopory, 1995, Ecker, 1995, Zhou et al., 1995). The role of protein phosphorylation in low temperature gene regulation is inferred from the low temperature inducibility of the numerous kinases already mentioned as well as from cold-induced changes in phosphorylation patterns seen in alfalfa (Monroy et al., 1993b).

Two principle approaches have been used in studying the molecular biology of signal transduction pathways. One can study a promoter and its elements, determine what acts upon those promoter elements, what interacts with that factor and so on backward through the pathway. Alternatively, mutations can be induced and plants that show aberrant expression of the gene or genes of interest can be characterized. This approach can reveal members of the pathway in any order. The first approach in studying the pathway regulating *COR*15a has already borne some fruit. A promoter element necessary and sufficient for low temperature induction of *COR*78, the DRE or dehydration-responsive element, has been defined (Yamaguchi-Shinozaki and Shinozaki, 1994) and an element containing the core CCGAC of the DRE is found in the region believed to be responsible for cold-regulation of *COR*15a (Baker et al., 1994). In *Brassica napus* this same DRE-like element, CCGAC, has been shown by transient expression studies to be required for cold induction of *BN115* (Jiang et al., 1996).

Furthermore, a protein capable of binding to this element has been isolated from *Arabidopsis* (Stockinger et al., in preparation). Whether this protein is actually involved in the activation of *COR*15a or any of the other *COR* genes in vivo awaits further study.

Mutagenesis Strategies

When using mutants to examine low temperature signal transduction in a plant that can cold acclimate, there are two main strategies. Plants can be mutagenized and screened or selected looking for increased freezing tolerance without acclimation or for frost-sensitivity despite acclimation or they can be screened or selected more directly for the induction or lack of induction of *COR* genes. The first method has the advantage of a readily observable phenotype, but may result in mutants in a number of processes unrelated or tangentially related to signal transduction.

Using gene expression as the phenotype of interest requires more setup, and could still yield mutants not directly involved in the pathway, but is a more direct approach to the question. In order to identify mutants in the regulation of a gene whose expression has no obvious morphological phenotype, the promoter of the gene of interest is fused to the coding region of a reporter gene and plants

carrying this construct are mutagenized and screened or selected based on the expression of the reporter.

Variations of this approach have been used several times with varying results. Takahashi et al. (1992) fused the promoter of the *HSP18.2* gene to the gene encoding beta-glucuronidase (GUS) and screened using a 4-methyl umbelliferyl glucuronide (MUG) assay. Two of the mutants isolated showed a large reduction in GUS activity, but only a small reduction in the endogenous transcript, while the mutation in the third acted as a dominant trans-suppresser of the introduced gene without affecting the endogenous gene. When Brusslan et al. (1993) fused the *cab140* promoter to *tms2*, the protein product of which converts nontoxic naphthalene acetamide into toxic naphthalene acetic acid, their selection resulted in a mutant which had reduced levels of both the endogenous and the introduced gene, but not of any other phytochrome-regulated genes, nor could the phenotype be genetically separated from the T-DNA insert. From this they infer that the lowered levels of mRNA are a result of co-suppression caused by a mutation in the introduced gene.

Susek et al. (1993) employed two reporters in their construct, putting the CAB3 promoter in front of both the GUS gene and the gene for hygromycin resistance. They selected for hygromycin resistance and then checked the putative mutants for GUS activity to make sure the mutation wasn't in the transgene. This

resulted in the isolation of at least three genes necessary for coupling the expression of some nuclear genes to the functional state of the chloroplast. The Dong laboratory used the promoter of the *Arabidopsis* β-1,3-glucuronidase fused to the GUS coding region and their screen gave them both a mutant that is nonresponsive to inducers of systemic acquired resistance (SAR) (Cao et al., 1994) and one that leads to constitutive expression of SAR (Bowling et al., 1994).

While mutants were recovered with a screen and with the use of only one reporter, the development of a selection would allow a greater number of mutants to be tested with less time and energy, and the use of a second reporter does facilitate the identification of false positives. Given this information, it should be feasible to design and implement a mutagenesis approach for isolating genes involved in the signal trandsuction pathway responsible for the induction of cold-regulated genes.

Literature Cited

Adamska I, Kloppstech K (1994) Low temperature increases the abundance of early light-inducible transcript under light stress conditions. The Journ. of Biol. Chem. 269:30221-30226

Anderson JV, Li Q-B, Haskell DW, Guy CL (1994) Structural organization of the spinach endoplasmic reticulum-luminal 70-kilodalton heat-shock cognate gene and expression of 70-kilodalton heat-shock genes during cold acclimation. Plant Physiol. 104:1359-1370

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. 24: 701-713

Bowling SA, Guo A, Cao H, Gordon AS, Klessing DF, Dong X (1994) A mutation in *Arabidopsis* that leads to constitutive expression of systemic acquired resistance. The Plant Cell 6:1845-1857

Bramlage WJ, Meir S (1990) Chilling injury of crops of temperate origin. In, CY Wang, ed., Chilling injury of horticultural crops. CRC Press, Inc., Boca Raton, Florida, pp 37-49

Brusslan JA, Karlin-Neumann GA, Huang L, Tobin E (1993) An *Arabidopsis* mutant with a reduced level of *cab140* RNA is a result of cosuppression. The Plant Cell 5:667-677

Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. The Plant Cell 6:1583-1592

Castonguay Y, Laberge S, Nadeau P, Vezina LP (1994) A cold-induced gene from *Medicago sativa* encodes a bimodular protein similar to developmentally regulated proteins. Plant Mol. Biol. 24:799-804

Chen THH, Gavinlertvatana P, Li PH (1979) Cold acclimation of stem-cultured plants and leaf callus of *Solanum* species. Bot. Gaz. 140:142-147

Chen THH, Gusta LV (1983) abscisic acid-induced freezing resistance in cultured plant cells. Plant Physiol. 73:71-75

Chen H-H, Li PH, Brenner ML (1983) Involvement of abscisic acid in potato cold acclimation. Plant Physiol. 71:362-365

Chu B, Sunstad DP, Carter JV (1993) Alteration of β-tubulin gene expression during low-temperature exposure in leaves of *Arabidopsis thaliana*. Plant Physiol. 103:371-377

Colorado P, Rodriguez A, Nicolas G, Rodriguez D (1994) Abscisic acid and stress regulate gene expression during germination of chick pea seeds. Possible role of calcium. Physiol. Plant. 91:461-467

Cox W, Levitt J (1976) Interrelations between environmental factors and freezing resistance of cabbage leaves. Plant Physiol. 57:553-555

Crespi MD, Zabaleta EJ, Pontis HG, Salerno GL (1991) Sucrose synthetase expression during cold-acclimation in wheat. Plant Physiol. 96:887-891

Crosatti C, Soncici C, Stanca AM, Cattiveli L (1995) The accumulation of a cold-regulated chloroplastic protein is light-dependent. Planta 196:458-463

Ding JP, Pickard BG (1993) Modulation of mechanosensitive calcium-selective cation channels by temperature. The Plant Journ. 3:713-720

Dowgert MF, Steponkus PL (1983) Effect of cold acclimation on intracellular ice formation in isolated protoplasts. Plant Physiol. 72:978-988

Dunn MA, Goddard NJ, Zhang L, Pearce RS, Hughes MA (1994) Low-temperature-responsive barley genes have different control mechanisms. Plant Mol. Biol. 24:879-888

Dunn MA, Morris A, Jack PL, Hughes MA (1993) A low-temperature-responsive translation elongation factor 1α from barley (*Hordeum vulgare* L.). Plant Mol. Biol. 23:221-225

Ecker JR (1995) The ethylene signal transduction pathway in plants. Science 268:667-675

Gilmour SJ, Lin C, Thomashow, MF (1996) Purification and properties of *Arabidopsis thaliana COR* (cold-regulated) gene polypeptides COR15am and COR6.6 expressed in *Escherichia coli*. Plant Physiol. 111:293-299

Gilmour SJ, Hajela RK, Thomashow MF (1988) Cold acclimation in *Arabidopsis thaliana*. Plant Physiol. 87:745-750

Gilmour SJ, Thomashow MF (1991) Cold acclimation and cold-regulated gene expression in ABA mutants of *Arabidopsis thaliana*. Plant Mol. Biol. 17:1233-1240

Guo W, Ward RW, Thomashow MF (1992) Characterization of a cold-regulated wheat gene related to *Arabidopsis cor47*. Plant Physiol. 100:915-922

Guy C (1990) Cold acclimation and freezing stress tolerance: role of protein metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41:187-23

Guy CL, Haskell D (1988) Detection of polypeptides associated with the cold acclimation process in spinach. Electrophoresis 9:787-796

Guy CL, Haskell D, Neven L, Klein P, Smelser C (1992) Hydration-stable responsive protein link cold and drought stress in spinach. Planta 188:265-270

Hajela RK, Horvath DP, Gilmour SJ, Thomashow MF (1990) Molecular cloning and expression of *cor* (Cold-regulated) genes in *Arabidopsis thaliana*. Plant Pysiol. 93:1246-1252

Holappa LD, Walker-Simmons MK (1995) The wheat abscisic acid-responsive protein kinase mRNA, PKABA1, is up-regulated by dehydration, cold temperature, and osmotic stress. Plant Physiol. 108:1203-1210

Hon W-C, Griffith M, Mlynarz A, Kwok YC, Yang DSC (1995) Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. Plant Physiol. 109:879-889

Horvath DP, McLarney BK, Thomashow MF (1993) Regulation of *Arabidopsis thaliana* L. (Heyn) *cor78* in response to low temperature. Plant Physiol. 103:1047-1053

Houde MJ, Laliberte J-F, Rassart E, Dhindsa RS, Sarhan F (1992) Cloning, characterization, and expression of a cDNA encoding a 50-kilodalton protein specifically induced by cold acclimation in wheat. Plant Physiol. 99:1381-1387

Hughes MA, Dunn MA, Pearce RS, White AJ, Zhang L (1992) An abscisic-acid-responsive, low temperature barley gene has homology with a maize phospholipid transfer protein. Plant, Cell and Environ. 15:861-865

Hugly S, Somerville C (1992) A role for membrane lipid polyunsaturation in chloroplast biogenesis at low temperature. Plant. Physiol. 99:197-202

Jarillo JA, Capel J, Leyva A, Martinez-Zapater JM, Salinas J (1994) Two related low-temperature-inducible genes of *Arabidopsis* encode proteins showing high homology to 14-3-3 proteins, a family of putative kinase regulators. Plant Mol. Biol. 25:693-704

Jarillo JA, Levya A, Salinas J, Martinez-Zapater JM (1993) Low temperature induces the accumulation of alcohol dehydrogenase mRNA in *Arabidopsis thaliana*, a chilling-tolerant plant. Plant Physiol. 101:833-837

Jiang C, Iu B, Singh J (1996) Requirement of a CCGAC *cis*-acting element for cold induction of the *BN115* gene from winter *Brassica napus*. Plant Mol. Biol. 30:679-684

Kendall EJ, McKersie BD (1989) Free radical and freezing injury to cell membranes of winter wheat. Physiol. Plant. 76:86-94

Knight MR, Campbell AK, Smith SM, Trewavas AJ (1991) Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. Nature 352:524-526

Knight H, Trewavas AJ, Knight MR (1996) Cold calcium signaling in Arabidopsis involves two cellular pools and a change in calcium signature after acclimation. The Plant Cell 8:489-503

Krishna P, Sacco M, Cherutti JF, Hill S (1995) Cold-induced accumulation of hsp90 transcripts in *Brassica napus*. Plant Physiol. 107:915-923

Kurkela S, Borg-Franck M (1992) Structure and expression of *kin2*, one of two cold- and ABA-induced genes of *Arabidopsis thaliana*. Plant Mol. Biol. 19:689-692

Kurkela S, Franck M (1990) Cloning and characterization of a cold- and ABA-inducible Arabidopsis gene. Plant Mol. Biol. 15:137-144

Lalk I, Dorffling K (1985) Hardening, abscisic acid, proline and freezing resistance in two winter wheat varieties. Physiol. Plant. 63:287-292

Lang V, Heino P, Palva ET (1989) Low temperature acclimation and treatment with exogenous abscisic acid induce common polypeptides in *Arabidopsis thaliana* (L.) Heynh. Theor. Appl. Genet. 77:729-734

Lang V, Palva ET (1992) The expression of *rab*-related gene, *rab18*, is induced by abscisic acid during the cold acclimation process of *Arabidopsis thaliana* (L.) Heynh. Plant Mol. Biol. 20:951-962

Lee SP, Chen THH (1993) Molecular cloning of abscisic acid-responsive mRNAs expressed during the induction of freezing tolerance in bromegrass (*Bromus inermis* Leyss) suspension culture. Plant Physiol. 101:1089-1096

Levya A, Jarillo JA, Salinas J, Martinez-Zapater JM (1995) Low temperature induces the accumulation of *phenylalanine ammonia-lyase* and *chalcone synthase* mRNAs of *Arabidopsis thaliana* in a light-dependent manner. Plant Physiol. 108:39-46

Li Q-B, Anderson JV, Guy CL (1994) A cDNA clone encoding a spinach 70-kilodalton heat-shock cognate. Plant Physiol. 105:457-458

Lin C, Thomashow MF (1992) DNA sequence analysis of a complementary DNA for cold-regulated *Arabidopsis gene cor15* and characterization of the COR15 polypeptide. Plant Physiol. 99:519-525

Los D, Horvath I, Vigh L, Murata N (1993) The temperature-dependent expression of the desaturase gene *desA* in *Synechosystis* PCC6803. FEBS LETTERS 318:57-60

Lu G, deLisle AJ, de Vetten NC, Ferl RJ (1992) Brain proteins in plants: An *Arabidopsis* homolog to neurotransmitter pathway activators is part of a DNA binding complex. PNAS 89:11490-11494

Lynch DV, Steponkus PL (1987) Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). Plant Physiol. 83:762-767

Markhart III AH (1986) Chilling injury: a review of possible causes. Hortscience 21:1329-1333

Mantyla E, Lang V, Palva ET (1995) Role of abscisic acid in drought-induced freezing tolerance, cold acclimation, and accumulation of LTI78 and RAB18 proteins in *Arabidopsis thaliana*. Plant Physiol. 107:141-148

Marana C, Garcia-Olmedo F, Carbonero P (1990) Differential expression of two types of sucrose synthetase-encoding genes in wheat in response to anaerobiosis, cold shock and light. Gene 88:167-172

McKersie BD, Chen Y, de Beus M, Bowley SR, Bowler C, Inze D, D'Halluin K, Botterman J (1993) Superoxide dismutase enhances tolerance of freezing stress in transgenic alfalfa (*Medicago sativa* L.). Plant Physiol. 103:1155-1163

Minorsky PV (1985) An heuristic hypothesis of chilling injury in plants: a role for calcium as the primary physiological transducer of injury. Plant, Cell and Environ. 8:75-94

Minorsky PV (1989) Temperature sensing by plants: a review and hypothesis. Plant, Cell and Environ. 12:119-135

Miquel M, James D, Dooner H, Browse J (1993) *Arabidopsis* requires polyunsaturated lipids for low-temperature survival. PNAS 90:6208-6212

Mizoguchi T, Hayashida N, Yamaguchi-Shinozaki K, Kamada H, Shinozaki K (1995) Two genes that encode ribosomal-protein S6 kinase homologs are induced by cold or salinity stress in *Arabidopsis thaliana*. FEBS Lett. 358:199-204

Mizoguchi T, Irie K, Hirayama T, Hayashida N, Yamaguchi-Shinozaki K, Matsumoto K, Shinozaki K (1996) A gene encoding a mitogen-activated protein kinase kinase kinase is induced simultaneously with genes for a mitogen-activated kinase and an S6 ribosomal protein kinase by touch, cold and water stress in *Arabidopsis thaliana*. PNAS 93:765-769

Mohapatra SS, Wolfraim L, Poole RJ, Dhindsa RS (1989) Molecular cloning and relationship to freezing tolerance of cold-acclimation-specific genes of alfalfa. Plant Physiol. 89:375-380

Monroy AF, Castonguay Y, Laberge S, Sarhan F, Vezina LP, Dhindsa RS (1993a) A new cold-induced alfalfa gene is associated with enhanced hardening at subzero temperature. Plant Physiol. 102:873-879

Monroy AF, Sarhan F, Dhindsa RS (1993b) Cold-induced changes in freezing tolerance, protein phosphorylation, and gene expression. Plant Physiol. 102:1227-1235

Monroy AF, Dhindsa RS (1995) Low-temperature signal transduction: Induction of cold acclimation-specific genes of alfalfa by calcium at 25°C. The Plant Cell 7:321-331

Moynihan MR, Ordentlich A, Raskin I (1995) Chilling-induced heat evolution in plants. Plant Physiol. 108:995-999

Murata T (1990) Relation of chilling stress to membrane permeability. In, CY Wang, ed., Chilling injury of horticultural crops. CRC Press, Inc., Boca Raton, Florida, pp 201-209

Murata N, Nishida I (1990) Lipids in relation to chilling sensitivity in plants. In, CY Wang, ed., Chilling injury of horticultural crops. CRC Press, Inc., Boca Raton, Florida, pp 181-199

Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y, Nishida I (1992) Genetically engineered alteration in the chilling sensitivity of plants. Nature 356:710-713

Neven LG Haskell DW, Guy CL, Denslow N, Klein PA, Green LG, Silverman A (1992) Association of 70-kilodalton heat-shock cognate proteins with acclimation to cold. Plant Physiol. 99:1362-1369

Neven LG, Haskell DW, Hofig A, Li Q-B, Guy CL (1993) Characterization of a spinach gene responsive to low temperature and water stress. Plant Mol. Biol. 21:291-305

Newsted WJ, Chibbar RN, Georges F (1991) Effect of low temperature stress on the expression of sucrose synthetase in spring and winter wheat plants. Development of a monoclonal antibody against wheat germ sucrose synthetase. Biochem. Cell Biol. 69:36-41

Nordin K, Vahala T, Palva ET (1993) Differential expression of two related, low-temperature-induced genes in *Arabidopsis thaliana* (L.) Heynh. Plant Mol. Biol. 21:641-653

Orr W, White TC, Iu B, Robert L, Singh J (1995) Characterization of a low-temperature-induced cDNA from winter *Brassica napus* encoding the 70kDa subunit of tonoplast ATPase. Plant Mol. Biol. 28:943-948

Patterson BD, Reid MS (1990) Genetic and environmental influences on the expression of chilling injury. In, CY Wang, ed., Chilling injury of horticultural crops. CRC Press, Inc., Boca Raton, Florida, pp 87-112

Pihakaski-Maunsbach K, Puhakainen T (1995) Effect of cold exposure on cortical microtubules of rye (*Secale cereale*) as observed by immunocytochemistry. Physiol. Plant. 93:563-571

Purvis AC, Shewfelt RL (1993) Does the alternate pathway ameliorate chilling injury in sensitive plant tissues? Phys. Plant. 88:712-718

Raghuram N, Sopory SK (1995) Evidence for some common signal transduction events for opposite regulation of nitrate reductase and phytochrome-I gene expression by light. Plant Mol. Biol. 29:25-35

Raison JK, Orr GR (1990) Proposals for a better understanding of the molecular basis of chilling injury. In, CY Wang, ed., Chilling injury of horticultural crops. CRC Press, Inc., Boca Raton, Florida, pp 145-164

Saez-Vasquez J, Raynal M, Delsney M (1995) A rapeseed cold-inducible transcript encodes a phosphoenolpyruvate carboxykinase. Plant Physiol. 109:611-618

Sakai A, Larcher W (1987) Frost survival of plants: response and adaptation to freezing stress. Spring-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo

Sakai A, Sugawara Y (1978) Survival of plant germplasm in liquid nitrogen. In, PH Li, A Sakai, eds., Plant cold hardiness and freezing stress: mechanisms and crop implications. Academic Press, New York, San Francisco, London, pp 345-359

Siminovitch D, Cloutier Y (1993) Drought and freezing tolerance and adaptation in plants: Some evidence for near equivalences. Cryobiology 20:487-503

Steponkus PL (1984) Role of the plasma membrane in freezing injury and cold acclimation. Ann. Rev. Plant Physiol. 35:543-584

Steponkus PL, Lynch DV (1989) Freeze/thaw-induced destabilization of the plasma membrane and effects of cold acclimation. J. of Bioenergetics and Biomembranes 21:21-41

Susek RE, Ausubel FM, Chory J (1993) Signal transduction mutants of *Arabidopsis* uncouple nuclear *CAB* and *RBCS* gene expression from chloroplast development. Cell 74:787-799

Sutton F, Ding X, Kenefick DG (1992) Group 3 LEA gene HVA1 regulation by cold acclimation and deacclimation in two barley cultivars with varying freeze resistance. Plant Physiol. 99:338-340

Takahashi T, Naito S, Komeda Y (1992) The *Arabidopsis HSP18.2* promoter/GUS gene fusion in transgenic *Arabidopsis* plants: a powerful tool for the isolation of regulatory mutants of the heat-shock response. The Plant Journ. 2:751-761

Thomashow MF (1993) Genes induced during cold acclimation in higher plants. In, PL Steponkus, ed., Advances in low temperature biology, vol. 2. JAI Press Ltd., London, pp 183-210

Thomashow MF (1994) *Arabidopsis* as a model for studying mechanisms of plant cold tolerance. In, EM Meyerowitz, CR Somerville, eds., *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 807-834

Uemura M, Joseph RA, Steponkus PL (1995) Cold acclimation *of Arabidopsis thaliana*: Effect of plasma membrane lipid composition and freeze-induced lesions. Plant Physiol. 109:15-30

van Zee K, Chen FQ, Hayes PM, Close TJ, Chen THH (1995) Cold-specific induction of a dehydrin gene family member in barley. Plant Physiol. 108:1233-1239

Vigh L, Los DA, Horvath I, Murata N (1993) The primary signal in the biological perception of temperature: Pd-catalyzed hydrogenation of membrane lipids stimulated the expression of the *desA* gene in *Synechocystis* PCC 6803. PNAS 90:9090-9094

Wang X, Sato R, Brown MS, Hua X, Goldstein JL (1994) *SREBP-1*, a membrane-bound transcription factor released by sterol-regulated proteolysis. Cell 77:53-62

Webb MS, Steponkus PL (1993) Freeze-induced membrane ultrastructural alterations in rye (*Secale cereale*) leaves. Plant Physiol. 101:955-963

Weiser RL, Wallner SJ, Waddell JW (1990) Cell wall and extensin changes during cold acclimation of pea seedlings. Plant Physiol. 93:1021-1026

Welin BV, Olson A, Nylander M, Palva ET (1994) Characterization and differential expression of *dhn/lea/rab*-like genes during cold acclimation and drought stress in *Arabidopsis thaliana*. Plant Mol. Biol. 26:131-144

Welin BV, Olson A, Palva ET (1995) Structure and organization of two closely related low-temperature-induced *dhn/lea/rab*-like genes in *Arabidopsis thaliana* L. Heynh. Plant Mol. Biol. 29:391-395

White AJ, Dunn MA, Brown K, Hughes MA (1994) Comparative analysis of genomic sequence and expression of a lipid transfer protein gene family in winter barley. Journ. of Exp. Bot. 45:1885:1892

Wilhelm KS, Thomashow MF (1993) *Arabidopsis thaliana COR15b*, an apparent homologue of *COR15a*, is strongly responsive to cold and ABA, but not drought. Plant Mol. Biol. 23:1073-1077

Williams J, Bulman M, Huttly A, Phillips A, Neill S (1994) Characterization of a cDNA from *Arabidopsis thaliana* encoding a potential thiol protease whose expression is induced independently by wilting and abscisic acid. Plant Mol. Biol. 25:259-270

Wilson JM (1979) Drought resistance as related to low temperature stress. In, JM Lyons, D Graham, JK Raison, eds., Low temperature stress in crop plants: the role of the membrane. Academic Press, New York, London, Sydney, Toronto, San Francisco, pp 47-65

Wolfraim LA, Langis R, Tyson H, Dhindsa RS (1993) cDNA sequence, expression and transcript stability of a cold acclimation-specific gene, *cas18*, of alfalfa (*Medicago falcata*) cells. Plant Physiol. 101:1275-1282

Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of the expression of a desiccation-responsive *rd29* gene of *Arabidopsis thaliana* and the analysis of its promoter in transgenic plants. Mol. Gen. Genet. 236:331-340

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. The Plant Cell 6:251-264

Zhou J, Loh Y-T, Bressan RA, Martin GB (1995) The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. Cell 83:925-935

Chapter 2:

Screen for mutations in the signal transduction pathway responsible for induction of *COR*15a

Summary

In preparation for gene regulation studies at the promoter level, a genomic clone containing *COR*15a, a gene inducible by low temperature, drought and abscisic acid (ABA) and encoding a chloroplast-targeted protein, was isolated. Approximately one kilobase of the *COR*15a promoter was then taken from this clone and fused to the coding region of beta-glucuronidase (GUS) in order to find elements of the signal transduction pathway responsible for the induction of *COR*15a. Transgenic plants carrying this construct were mutagenized and screened for the altered expression of GUS activity in warm-grown or cold-treated plants. Plants detectably altered in GUS expression were found in both screens. Some of these plants do not contain the transgene, while others appear to have a mutation that only affects the transgene. None show altered expression of endogenous *COR*15a, and hence do not carry mutations in the signal transduction pathway responsible for the induction of *COR*15a. Means of improving the mutagenesis strategy are discussed.

Introduction

Plants, being stationary poikilotherms, require that a series of physiological changes occur when the temperature becomes too low for optimal growth. Of those plants that are able to withstand chilling temperatures, some are also able to achieve an increased level of freezing tolerance after exposure to low non-freezing temperatures in a process known as cold acclimation (Levitt, 1980). Among the many metabolic alterations that occur during this process are changes in gene expression (Thomashow, 1994).

In *Arabidopsis thaliana*, a plant that is able to cold acclimate (Gilmour et al., 1988), the mRNA levels of a number of *COR* (*co*ld-*r*egulated) genes have been found to increase dramatically during low temperature treatment (Thomashow, 1993). A subset of the genes induced at low temperature, including members of gene families *COR*78, *COR*47, *COR*15 and *COR*6.6 have a number of properties in common. The message levels of these genes remain high as long as plants are maintained at low temperature, but decrease rapidly when the plants are returned to normal growth temperatures (Thomashow, 1994). They are also induced by water stress and/or by the exogenous application of abscisic acid (ABA), but not by heat shock. Finally, the proteins encoded by these genes remain soluble upon boiling, a property that is thought to be a consequence of their high hydrophilicity (Hajela et al., 1990).

The mechanism by which *COR* genes are regulated is of interest for practical as well as purely academic reasons. A program using genetic engineering to alter plant freezing tolerance must take into account the regulation of the genes involved as well as the properties of the genes themselves. At the start of this work, nothing was known about *cis*- or *trans*-acting elements involved in the induction of any member of the four *COR* gene families. When experiments analyzing the regulation of the *COR* genes were initiated, nuclear run-on assays indicated that only the *COR*15 gene family was regulated primarily at the transcriptional level (Hajela et al., 1990). Thus, it was chosen for promoter analysis experiments.

A genomic clone containing *COR*15a was isolated and promoter deletion analysis was performed. In this analysis, the *COR*15a promoter was found to be inducible by low temperature, drought and exogenously applied ABA (Baker et al., 1994). A mutagenesis approach designed to identify proteins involved in a signal transduction pathway that activates transcription of *COR*15a was also undertaken. Because there is no easily detectable phenotype specific for the induction of *COR*15a, one of the constructs used in the promoter deletion analysis, which contained approximately one kilobase of the *COR*15a promoter fused to the GUS coding region, was used as a screenable marker. The expression of GUS from this construct should mimic expression of the endogenous *COR*15a gene in transgenic plants.

Mutational analyses of reporter gene fusions of this type have been successfully used to examine the signals responsible for the induction of genes responsive to heat-shock (Takahashi et al., 1992), to systemic acquired resistance (Bowling et al., 1994 and Cao et al., 1994), and to chloroplast development (Susek et al., 1993) although the identities of the mutated genes have yet to be reported. A screen using the *cab140* promoter fused to *tms2* yielded, unexpectedly, a cosuppressed mutant (Bruslan et al. 1993).

It was hoped that mutants obtained could be used to determine whether all the *COR* genes were controlled by a common pathway, to see if mutants altered in their response to one of the inducing stimuli would also be altered in their response to the others, to estimate how many steps were in the pathway and ultimately to clone the genes involved. In the end, however, the experiment only suggested ways to improve the screen, as no plants with mutations in the signal transduction pathway involved in the induction of *COR*15a were identified.

Results

Isolation of a genomic clone encoding *COR*15a

A genomic library of *Arabidopsis* DNA in λEMBL3 was screened using pHH67 (Hajela, 1990), a cDNA corresponding to *COR*15a, as a probe. A phage containing an insert of approximately 12 kilobases of DNA, with the 5' end of *COR*15a located approximately 3 kilobases from one end, was isolated. The insert DNA from this phage was subcloned into pUC19 and was further manipulated in that form.

Deletion studies (Baker et al., 1994) demonstrated that the *COR*15a promoter is responsive to low temperature, drought and exogenously applied abscisic acid (ABA).

Mutagenesis strategy

Because there is no recognized phenotype for the induction of *COR*15a, a system was devised to facilitate a screen for mutants that were altered in the signal transduction pathway responsible for the induction of this gene (Figure 2.1). The promoter of *COR*15a fused to the coding region of GUS was introduced into plants for use as a screenable marker. Transgenic plants that expressed the reporter gene, GUS, at a cold temperature (4°C), but not at normal growth

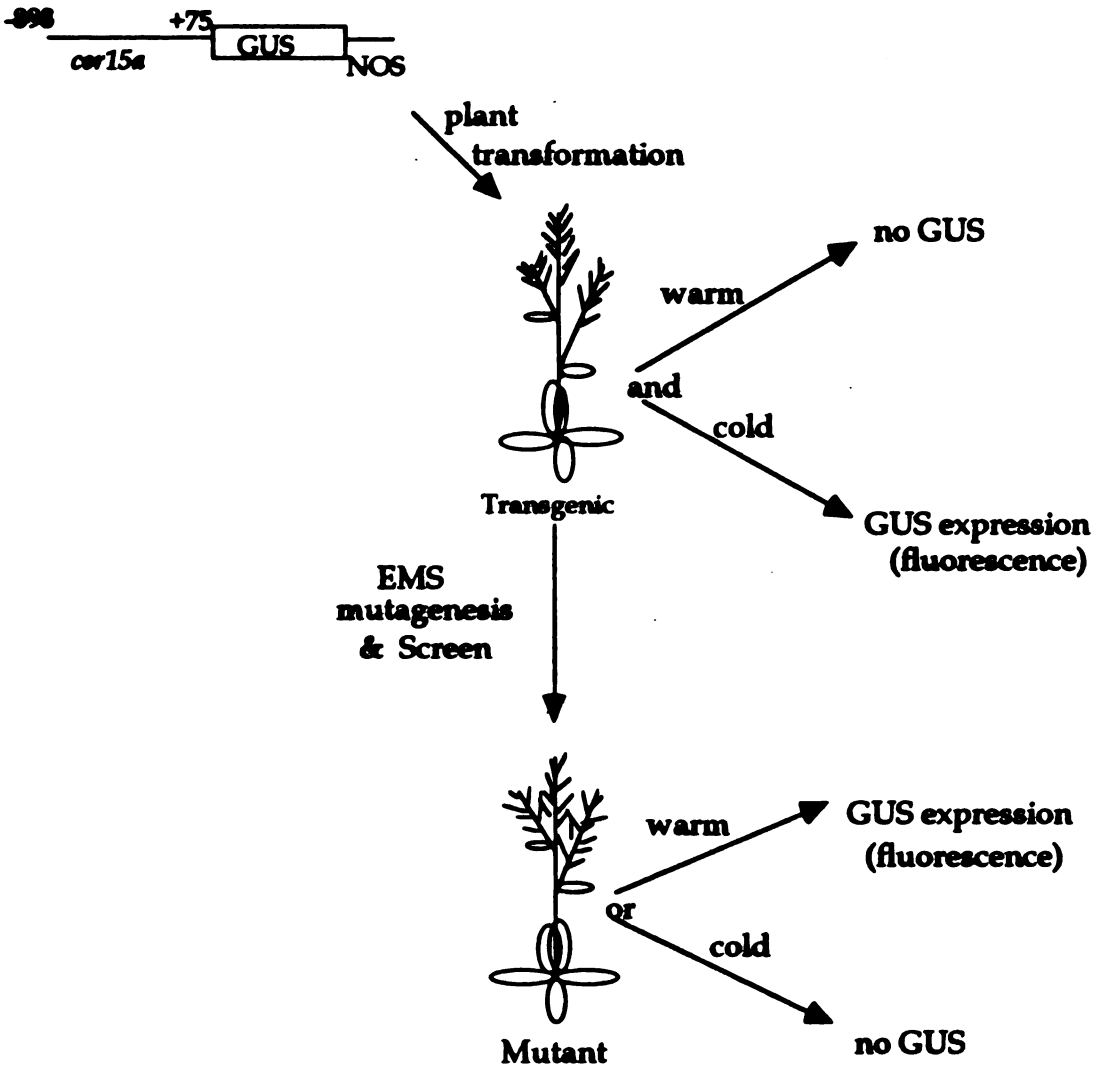


Figure 2.1 Screen for signal transduction mutants. A construct consisting of the promoter of *COR*15a, the coding region of GUS and a NOS terminator was introduced into *Arabidopsis*. A transgenic line that strongly expressed GUS at low temperature but not at normal growth temperatures and was homozygous for the transgenic insert was EMS mutagenized and screened using a MUG assay. Plants kept at normal growth temperatures were then screened for GUS expression and cold-treated plants were tested for the lack of GUS expression. Contrary to depiction, plants were screened after about three weeks of growth, at which point they were not yet flowering.

temperature (22°C), were EMS (ethyl methane sulfonate) mutagenized and screened at 22°C, looking for those with elevated expression, or at 4°C, looking for decreased expression.

Although cold induction is the response of interest, alteration of the expression of this promoter could also be due to a mutation in the signal(s) initiated by drought or ABA, since the *COR*15a promoter is also responsive to these stresses.

Preparation of transgenic reporter lines, mutagenesis and screening

A line of transgenic plants containing approximately one kilobase of the *COR*15a promoter (the -900/+78 fragment described by Baker *et al.*, 1994) fused to the GUS coding region of pBI 101.2 (Jefferson *et al.*, 1987) was selected because of its strong cold-inducible expression of GUS as tested by histochemical analysis. It has been shown that this construct is also inducible by drought and by exogenously applied ABA (Baker et al., 1994). The R3 generation (the third generation after regeneration of the transgenic plants) was tested for homozygosity first on Kanamycin plates and then by genomic Southerns done on eleven individual plants (Figure 2.2). Further genomic Southern analysis of this line (referred to as 1,6) suggests that there are three inserts at two loci and that the two linked inserts are inverted with respect to each other (Figure 2.3). It was hoped that having more than one insert would prevent the isolation of *cis* knockout mutations in the transgene, since it should be less likely for multiple

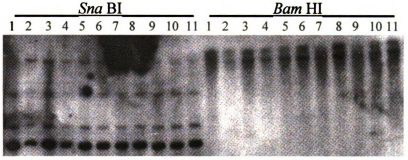


Figure 2.2 Test for homozygosity of inserts in line 1,6, using genomic Southerns. A line of transgenic plants containing a -900 to +78 *COR*15a promoter fragment fused to the GUS coding region was tested for homozygosity of the insert(s). Preparations of genomic DNA from 11 individual R3 plants of transgenic line 1,6 were digested with *Sna* BI, which cuts once within the inserted DNA, and with *Bam* HI, which does not cut within the insert, to determine if the previous generation (R2) of line 1,6 was homozygous for the transgenic insert.

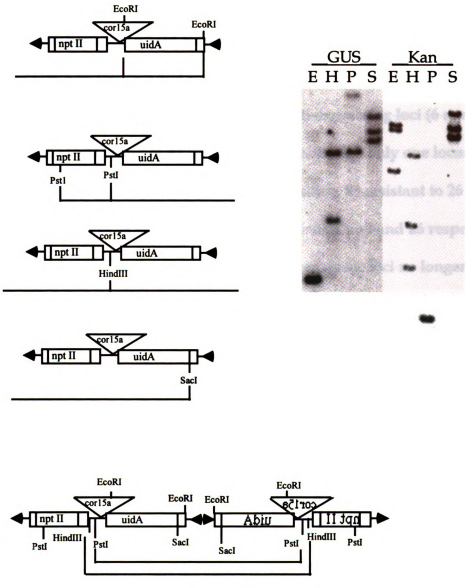


Figure 2.3 Putative organization of inserts in line 1,6. Genomic DNA isolated from transgenic line 1,6 was digested with enzymes *Eco*RI (E), *Hin*dIII (H), *Pst*I (P) and *Sac*I (S) and probed for the GUS (*uidA*) or Kan coding regions. The autoradiograms are shown to the right and the putative genomic arrangement is diagrammed at the left. The three bands in the *Sac*I digest suggest that there are three copies of the insert, and the GUS band shared between the *Hind* III and *Pst* I digests could be due to two copies of the transgene inserted tail to tail. The insert thought to be separate from the two linked inserts is shown four times. once for each restriction enzyme. All four enzymes are shown at once on the diagram of the two adjoining inserts.

inserts to be mutated at once. However, it also provides more targets for *cis*-acting mutations that might increase expression in the warm. Genetic analysis of a cross between this line and nontransgenic *Arabidopsis* (ecotype Columbia) gave ratios consistent with the presence of two GUS-expressing loci (6 noninducible plants out of 100 tested in F2 population a6), although only one locus conferred Kanamycin resistance (120 resistant to 41 sensitive, 83 resistant to 26 sensitive and 62 resistant to 36 sensitive, in F2 populations a2, a3 and a6 respectively). This suggests that the insert(s) at one of the transgenic loci no longer has a functional gene for kanamycin resistance.

Seeds of line 1,6 were then mutagenized with EMS and planted in 12 flats. Insect damage (principally due to thrips and aphids) severely limited the M2 (second mutant generation) seed production of this and of the subsequent round of mutagenesis. M2 seed was harvested in bulk from each flat, resulting in 12 pools of seed. An estimated 6,560 seeds were planted on plates containing DAP (2,6-diaminopurine) to test the effectiveness of mutagenesis. Plants homozygous for recessive mutations in the *apt* gene are able to grow in the presence of DAP, so the number of resistant plants provides a measure of the effectiveness of mutagenesis (Moffatt and Somerville, 1988). Four resistant plants were observed, demonstrating that the phenotype of a single recessive mutation was detectable in about one out of 1,600 plants. Later, more seed was mutagenized

in the same manner and 12 more pools were generated. The DAP test was not performed on this mutagenized seed.

Seed from each pool was planted in pots that were divided into quarters. Leaves of individual plants from each quadrant were clipped and placed into a 2 mM methylumbelliferyl glucuronide (MUG) solution in microtiter plate wells (Figure 2.4). Quadrants that contained a leaf of mutant phenotype (GUS expression at normal growth temperatures or a lack of GUS expression 4°C) were then rescreened and each plant was marked with a numbered toothpick. About half of the pots of plants were screened in the cold room and the other half in the growth chamber at standard growth conditions.

Putative mutants recovered

In the first round of screening, approximately 26 strong and 54 weak putative mutants in the warm screen and 36 strong and 67 weak putative mutants in the cold screen were selected and allowed to set seed. The next generation (M3) of the strong putative mutants was tested. Six of those selected in the warm screen (6-8#1w, 6-8#4w, 6-8#10w, 6-8#15w, 18#6w and 13-7#1w) and one of those selected in the cold screen (3-2#4c) again showed the mutant phenotype. The weak putative mutants were also tested, but none of those isolated in the warm screen displayed the mutant phenotype in the second generation. The twelve putative mutants from the cold screen that still showed a weak mutant

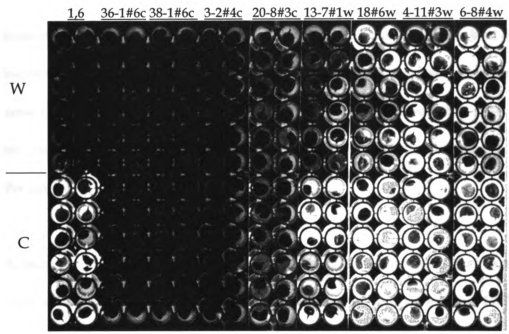


Figure 2.4 MUG assay of putative mutants. Two columns of microtiter plate wells are shown for each line of plants, with nonmutagenized 1,6 serving as the control. Leaves in the wells in the top half of the plate were clipped from warm-grown plants (W, 22°C), while those in the bottom half of the plate were taken from cold-grown plants (C, 2.5°C).

phenotype were not studied further because the variability of their phenotype made confirmation and characterization difficult. Initial testing of the non-mutagenized reporter transgenic had made it appear that the warm and cold response of the MUG assay was clear, but substantial variability was seen during screening, even in the controls, making it difficult to identify true mutants. The reason for this variability is not known, although it has been seen that different leaves from the same plant can give different levels of expression of the reporter gene (Figure 2.5). The number of false positives was decreased in later screening by retesting the putative mutants a week or two after the initial screen, although the improvement was not quantified.

A second round of screening turned up three putative mutants that expressed GUS at normal growth temperatures and three that did not express GUS at low temperature. Of these, two of those that express GUS at normal growth temperatures (6-10#14w and 4-11#3w) and one of those that does not express GUS at low temperature (20-8#3c) were still positive in the next generation. Two more plants that did not express GUS at low temperature (36-1c and 38-1c) were confirmed from yet a third round of screening. In all, approximately 13,000 plants were screened. Approximately 39 lines of putative mutants that do not express GUS at low temperature and 9 putative mutants that express GUS at normal growth temperatures remain to be tested at the M3 generation.

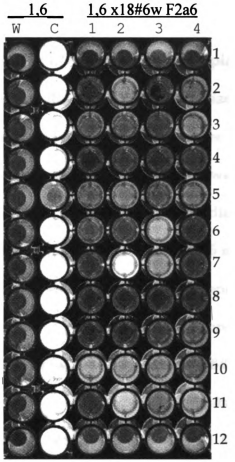


Figure 2.5 Variability of MUG assay. Four leaves from each of 12 warm-grown plants from the cross 1,6x18#6w (F2 generation, population a6) were screened for MUG activity. The identity of these plants is included for the sake of completeness, but is not important for the observation that different leaves from the same plant give different results in the MUG assay. Leaves from the twelve plants are shown in twelve rows, with the leaf number being indicated at the top of the microtiter plate. Leaf number does not correspond to the position of the leaf on the plant. Columns of nonmutagenized 1,6 grown at standard temperatures (W) and at 2.5°C (C) are included for comparison.

In order to determine whether lines which still showed the mutant phenotype in the M3 generation were affected in a signal transduction pathway responsible for inducing *COR*15a, northern blots of each were probed with GUS and *COR*15.

In all cases, the endogenous *COR*15 message of the plants selected as mutants was indistinguishable from that of the nonmutagenized transgenic control (Figure 2.6), except for line 13-7#1w, which was later found to contain a CaMV 35S promoter fragment (data not shown), and thus was not derived from the original transgenic line. Line 13-7#1w was not studied further. In the figure shown, it appears as though the endogenous *COR*15 message may be elevated over background for line 18#6w, but this was not reproducible. However, the GUS message was elevated in plants selected as expressing GUS at normal growth temperatures (4-11#3w, 6-8#1w, 6-8#4w, 6-8#10w, 6-8#15w, 6-10#14w, 13-7#1w and 18#6w) and was not detectable in plants chosen as not expressing GUS at low temperature (3-2#4c, 20-8#3c, 36-1c and 38-1c). In genomic Southern analysis, the latter were found to not contain the reporter (Figure 2.7), so they were not studied further. It seems unlikely that all three inserts of the transgene would have been deleted by EMS mutagenesis, suggesting that these may be contaminants. Indeed, a subsequent genomic Southern of line 38-1c showed that it contained an extra copy of the *COR*15a coding region indistinguishable from a CaMV 35S/*COR*15a sense transgenic used in the lab (data not shown).

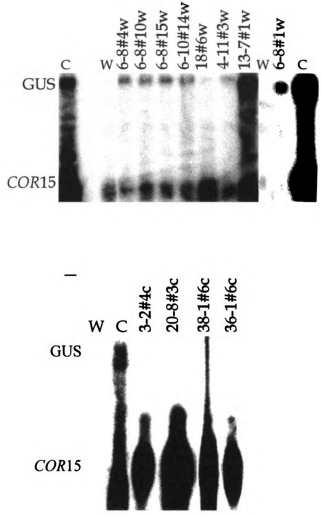


Figure 2.6 Expression of GUS and of *COR*15 mRNA in putative mutants. Samples of warm- (W) and cold- (C) grown nonmutagenized 1,6 are included as controls on each blot. Lines chosen as having increased GUS activity at 22°C (normal growth temperature) were grown at 22°C for this experiment. These are shown at the top of the page. Lines which did not have GUS activity at low temperatures (2.5°C) were grown at 2.5°C for this experiment and are shown at the bottom.

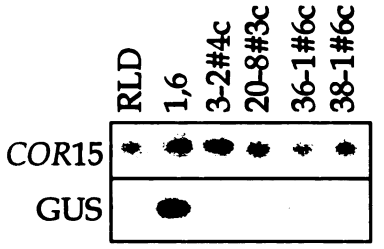


Figure 2.7 Southern analysis of putative mutants screened at low temperature. Duplicate blots of *Eco*RI-cut genomic DNA were probed for *COR*15 and for GUS sequences. RLD is an untransformed control, which should only contain *COR*15 sequences; 1,6 is the nonmutagenized transgenic control, which should contain GUS as well as COR15 sequences; and the other four lines are putative mutants that were chosen for their lack of GUS expression at low temperature. These were expected to contain both GUS and *COR*15 sequences. Both endogenous copies of *COR*15 are contained within a single *Eco*RI fragment, while the introduced copies of the transgene contain GUS on an *Eco*RI fragment internal to the insert (see Figure 2.3).

It was also noted that the GUS message produced by line 6-8#4w (selected as expressing GUS at normal growth temperatures) was still responsive to low temperature, drought and ABA treatment, and that the signal was lowered by incubation in water, as is true with the nonmutagenized control (Figure 2.8). Line 6-8#4w may be slightly less responsive to drought than 1,6. Lines 6-8#1w, 6-8#10w, 6-8#15w and 6-10#14w were not included in this or subsequent analyses because they are from the same pool of M2 seed as line 6-8#4w and are likely to be siblings of it. There has been no detected difference among these five putative mutants.

Genetic analysis of putative mutants

Because the mutations did not appear to be in the signal transduction pathway, and because *cis*-acting elements in the *COR*15a promoter are only partially understood, it was of interest to determine whether any of the mutations that caused elevated levels of GUS message in warm-grown plants were in the introduced gene. If so, and if the mutation was not in the GUS coding region, isolation and examination of the altered genes might give some insight into the function of *COR*15a promoter elements. If not, then the mutation would have to be in a gene that affected the introduced gene and not endogenous *COR*15, perhaps influencing the stability of the GUS message. Should such be the case, more care would have to be taken to avoid or quickly detect such mutations in subsequent mutagenesis attempts.

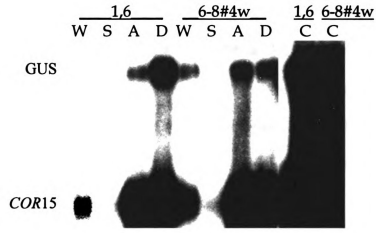


Figure 2.8 Inducibility of putative mutant line 6-8#4w. Plants of nonmutagenized 1,6 and putative mutant line 6-8#4w were grown under normal conditions (22°C, W), soaked in water (S), soaked in 100 μM ABA (A), drought-stressed (D) or cold-stressed (2.5°C, C). See materials and methods for details. The RNA was probed for GUS and for *COR*15 messages.

The first step in this analysis was to determine whether the mutant phenotype co-segregated with the introduced genes, and if so, with which locus. Lines 4-11#3w, 18#6w and 6-8#4w were crossed to the non-mutagenized reporter line and to nontransgenic *Arabidopsis*, ecotype Columbia. The F1 generation of each cross demonstrated the mutant GUS phenotype, and in the F2 generation, each putative mutant segregated at about 3 bright to 1 dark, indicating that the mutation was dominant. Southerns were done on individual bright plants of crosses between Columbia and lines 18#6w, 4-11#3w and 6-8#4w, and in each case, when there was only one locus represented, it was the one expected to contain the two linked transgenes (data shown for 4-11#3w in Figure 2.9). The eleven lines shown should be enough for 95% confidence of linkage (see Materials and Methods for details).

This co-segregation of the mutant phenotype with the transgenes was consistent with the idea that the mutation is either in the introduced gene or is linked to it, although nothing can be said regarding how tight the linkage is. It also supported the conclusion that the mutation is not in the signal transduction pathway and further suggested that it is also not in an unlinked gene that affects GUS stability.

Preliminary attempts to isolate the locus believed to contain the mutation were unsuccessful. Assuming that this locus does contain two copies of the transgene,

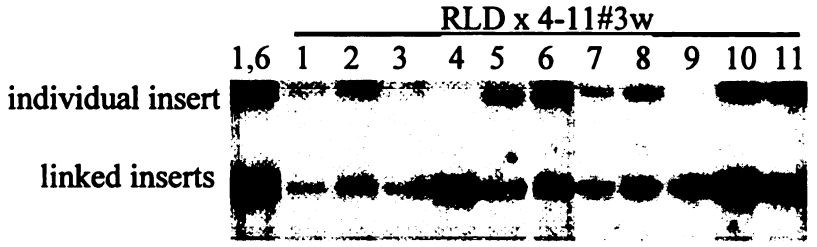


Figure 2.9 Cosegregation of the mutant phenotype with the locus thought to contain two linked copies of the reporter construct. Wildtype *Arabidopsis*, ecotype RLD, was crossed as the female parent to putative mutant line 4-11#3w. Genomic Southerns were performed on *Pst*I-cut DNA isolated from 11 individual F2 plants which had shown the mutant phenotype by the MUG assay. This digest (see Figure 2.3) separates two loci containing the transgene, with the upper band being that thought to contain a single insert and the lower band thought to contain two linked copies of the transgene. The blots were probed for the GUS coding region. Nonmutagenized line 1,6 is shown at left as the control.

both would need to be sequenced, and the possibility does still exist that the mutation is in the GUS coding region. For these reasons, it was determined that resources could be more effectively invested in less risky projects.

Test for polyploidy

One possible explanation for the lack of recessive mutations was that the original transgenic line may have become tetraploid during the tissue culture stage (Negrutiu et al., 1975) of the transformation process, thus making recessive mutations unobservable. However, if this were the case, the crosses of 1,6 to nontransgenic *Arabidopsis* should have produced aborted seed in the F2 generation because of the high amount of nondisjunction expected when a tetraploid and diploid are crossed. The F2 seed of this cross was wildtype in appearance and germinated normally, demonstrating that the transgenic was not tetraploid (data not shown).

Discussion

In the examination of *COR* gene regulation, a genomic clone was needed so that promoter analyses could be conducted. A cDNA corresponding to the coding region of *COR*15a was used to isolate from a genomic library a clone that contained about 3kb of sequence 5′ of *COR*15a and 8 to 10kb 3′. Deletion

analysis of the *COR*15a promoter found on this clone demonstrated that this promoter is inducible by low temperature, drought and ABA (Baker et al., 1994).

In an attempt to isolate mutants altered in the signal transduction pathway responsible for the induction of low temperature-responsive genes, plants containing the *COR*15a promoter fused to the GUS coding region were selected, EMS-mutagenized and screened for altered GUS expression. Approximately 6,000 plants were screened at normal growth temperatures and another 6,000 at 4°C and a large number of putative mutants was isolated. While these numbers are by no means exhaustive, growing mutant seed on DAP had already shown that one recessive mutation in the *apt* gene could be found in about 1,600 plants. In other systems, a screen looking for mutants in the signal transduction pathway for heat-shock inducible genes yielded about one mutant per 5,000 plants screened (Takahashi et al., 1992), and in other screens done using reporter constructs, one mutation was found in 2,000 or fewer plants screened (Cao et al., 1994, Bowling et al., 1994 and Brusslan et al., 1993). Practically, no more plants were screened because the available resources went into characterizing the mutants isolated, which turned out to be false positives.

Of those screened, only 8 (9.6%) of the initial putative mutants from the warm screen and 4 (3.7%) of the initial putative mutants from the cold screen still looked mutant in the next generation, indicating that the screen allowed for a

large percentage of false positives. The substantial variability seen during screening, even in the controls, may be due to the stability of the GUS protein (Jefferson, 1987) coupled with a sensitivity of the *COR*15a promoter to leaf to leaf variation in microenvironments. Testing a second leaf after the plant had a chance to grow for another week or two did eliminate a number of false positives more quickly than waiting to test the next generation.

Of the putative mutants that came through secondary screening, none were found to contain a mutation in the signal transduction pathway responsible for the induction of *COR*15a. Putative mutants screened at low temperature had either suffered deletion of all three copies of the transgene or, as was demonstrably true with line 38-1c, were contaminants, while those screened at normal growth temperatures only exhibited the mutant phenotype with regard to the transgene, and hence do not affect the signal transduction pathway of interest.

Since the mutant phenotype of the lines which express GUS at normal growth temperatures cosegregates with the locus thought to contain the two linked transgenes, the mutation must be either in one of the inserts or in a linked gene that affects the expression of GUS, but not of *COR*15 message. The former seems more likely, especially in line 6-8#4w, which shows the strongest phenotype and is still responsive to low temperature, drought, ABA and to suppression by

submergence in water. Lines 18#6w and 4-11#3w are still inducible by low temperature (data not shown), but the results from the water-soak test were less conclusive because the GUS expression of these lines is so weak at the normal growth temperature. This alteration in expression without loss of inducibility suggests that the mutation may be in the introduced *COR*15a promoter. Cloning this transgene from any of the three lines known to carry independent mutations, but particularly from line 6-8#4w, and determining what the mutation is might give insight into the function of *cis*-acting promoter elements. However, it has currently been determined that direct experiments fusing promoter elements to reporter genes are a more efficient way to study *cis*-acting elements, particularly because there is no way of knowing which of the two inserts thought to reside at this mutant locus might be responsible for the phenotype and because it is still possible that the mutation is within the coding region of GUS or is only linked to the locus.

There are a number of measures that could be taken in setting up a new attempt to find plants mutant in the signal transduction pathway responsible for inducing *COR*15a in order to improve the chances of finding the desired mutants. To reduce the amount of contamination, flowering plants should be kept in a separate chamber from those newly planted and, where possible, there should be an easily observable phenotype associated with the transgenic line used for mutagenesis which is unrelated to expression of *COR*15a or the

transgene so that contaminants can be recognized and discarded with a minimum amount of effort.

In order to make mutations in the transgene more readily identifiable, the original reporter construct could be redesigned to include a second reporter also fused to the *COR*15a promoter. This way, another relatively rapid test could be done on each putative mutant to see if the second reporter also demonstrated the mutant phenotype rather than waiting for a generation or two to get enough material for RNA isolation. Since mutations that appear to be in the locus containing the linked inserts were found in at least one out of 2,000 plants screened (counting the five mutants isolated from mutant pool 6 as a single mutation event), it might also be wise to select a transgenic line containing only one predictable insert, thus reducing the complexity of the system. The presence of the second reporter would obviate the need for a second insert to weed out null *cis*-mutations.

In light of the observed variability of GUS expression, it might also be wise to use the histochemical stain on several pot-grown plants of the transgenic lines proposed for use in mutagenesis to see if some lines show more variability from leaf to leaf than others.

Experiments in progress to tease apart the functions of various *cis*-acting elements may make it possible to design a construct that is only inducible by low temperature, thus making it more likely that the screen or selection will yield mutations related to low temperature induction and not to water relations. The promoter for *COR*15b may in fact only be responsive to low temperature (see Chapter 3).

Finally, the use of a selectable rather than a screenable reporter should make identification of mutants easier and should make the testing of a much larger number of plants possible. Resistance to an herbicide or antibiotic, such as BASTA (Bouchez et al., 1993), kanamycin (Valvekens et al., 1988) or hygromycin (Susek et al., 1993) could be used to select for plants that allow expression from the *COR*15a promoter under warm conditions and genes encoding proteins that convert a nontoxic substrate to a toxic substance, such as those for indoleacetamide hydrolase (Klee et al., 1987), nitrate reductase (Nussaume et al., 1991), or cytosine deaminase (Perera et al., 1993) could be used to select plants that do not have expression from the *COR*15a promoter at low temperature. For this low temperature screen, it would be advantageous to test the induction of the non-mutagenized reporter construct at varying temperatures, since the plants grow very slowly at 2°C and sensitivity to toxic compounds is often seen best after growth has occurred in their presence. It has been shown (see Chapter 4) that *COR*15a expression is increased above that in warm-grown controls even at

16°C, and that its expression is at or near its maximum at 8°C, so it should be possible to find a temperature at which expression is strong, and growth is not prohibitively slow.

It is possible that no mutants were found because the expression of *COR*15a is necessary during some stage of the life cycle of *Arabidopsis*. This view is consistent with the finding that staining was found in anthers of warm-grown plants containing the same construct as 1,6 (Baker et al., 1994). It was hoped that the use of EMS, rather than a mutagen that caused larger deletions, might allow some leaky mutations that would show alteration of expression without being lethal. Unfortunately, the variability of the screen led to focus on plants with strong phenotypes because the weaker ones were simply too hard to follow. Thus, steps taken to reduce the variability of the screen or selection take on added importance.

Further work could be directed toward rescreening the 39 lines of putative down in the cold mutants and 9 putative up in the warm mutants that remain to be tested at the M3 generation. This should, however, be secondary to the preparation of new reporter constructs for use in starting over because a new construct enabling a selection and the elimination of *cis* mutations should allow a more efficient identification of relevant mutations.

Materials and Methods

Library screening

The *A. thaliana* (ecotype Columbia) genomic library used to screen for *COR*15a genomic clones was a gift from Harry Klee (University of Florida). It was prepared by ligating a *Sau* 3AI partial digest of genomic DNA into the *Bam* HI site of lambda EMBL3 (Sambrook et al., 1989). The library was screened with pHH67 (Hajela et al., 1990), a cDNA representing *COR*15a, using standard methods (Sambrook et al., 1989). One of the recombinant phage that hybridized with pHH67, L67g, was characterized further. The *Arabidopsis* insert DNA was cut out as a *Sal* I fragment (*Sal* I sites flank the insert in the multiple cloning site of λEMBL3) and ligated into the *Sal* I site of pUC19 (Yanisch-Perron et al., 1985) to make plasmid p67gk. A 1.8 kb *Bgl* II fragment, which contains the coding region of *COR*15a and approximately 1 kb of 3' sequence, was subcloned into the *Bam* HI site of pBluescript (SK-) to make pBlue67.

Analysis of kanamycin resistance

Seeds were surface sterilized with 30% bleach/5% Triton X-100, washed with sterile water five times and plated on GM media with Kanamycin. GM media was made up of 0.05% 2-[N-morpholino]ethane sulfonic acid (MES), 0.8% phytagar and 1 bag/l of GibcoBRL Gamborg's B-5 Medium with sucrose. The

pH was adjusted to 5.7 with 1 M KOH and 50 mg/l Kanamycin was added after autoclaving. Seeds were plated in a sterile manner by drawing 100 µl of the seed/water mixture into a 1 ml pipetman tip, removing the tip from the pipetman and gently touching the tip to the surface of the agar to deposit seeds. Seeds were then separated using the same tip. The plates were wrapped with two pieces of Parafilm® and plants were grown at a light intensity of about 100 $\mu Es^{-1}m^{-2}$, a 16-hour photoperiod and temperatures of 23°C to 27°C for a week or until resistant plants (green) were differentiable from sensitive plants (white or yellow).

Histochemical analysis of GUS expression

Plate-grown seedlings were tested for cold-inducible expression of GUS using histochemical staining (Jefferson, 1987). Leaves of plants that had remained at growth conditions or that had been given an overnight treatment at 2°C were excised and incubated at 37°C overnight in a solution containing 100 mM sodium phosphate pH 7.0, 10 mM EDTA, 0.1% (v/v) Triton X-100, 2 mM 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (X-gluc), 1 mM $K_4Fe(CN)_6$ and 1 mM $K_3Fe(CN)_6$. Chlorophyll was removed by soaking the plant material overnight in 50% (v/v) EtOH. The tissue was then stored in 70% (v/v) EtOH.

Mutagenesis and screening

Mutagenesis was done essentially as described (Haughn and Somerville, 1986). For each round of mutagenesis, approximately 25,000 (500 mg) seeds of *Arabidopsis thaliana* L. (Heyn.) transgenic line 1,6 were placed in 100 mls of a 0.3% solution of EMS (ethyl methanesulfonate), left in a fume hood at room temperature for 8 hours and swirled occasionally. Seeds were then washed with reverse osmosis water 10 to 15 times, suspended in 0.1% phytagar and planted by pipeting the seed/agar suspension onto 21 flats of soil (about 1 seed per square cm). The M1 plants of the first round of mutagenesis were grown under fluorescent lights on a shelving unit at ambient conditions, while those of the second round were grown in a greenhouse. In both places, thrips and aphids infested the plants, limiting seed production. Safer™ insecticidal soap was used on the plants on the shelving unit and Dycarb and Orthene (applied by licensed personnel) were used on the plants in the greenhouse. The M2 seed that was produced was pooled by flat and planted out in pots that were divided into quadrants. Seedlings (approximately two weeks old) were screened for aberrant GUS expression. Plants screened in the warm were left in the chambers where they were grown, while those screened in the cold were moved to a 2°C cold room overnight initially, but for four days when it was recognized that longer cold treatment reduced the variability of the MUG screen. A scalpel was used to cut a leaf piece from each plant and forceps were used to place each piece in 250 µl of a solution containing 2 mM MUG (methylumbelliferyl glucuronide),

50 mM sodium phosphate pH 7.0 and 0.05% Triton X-100 in a microtiter plate well. These were then incubated at 37°C overnight the first few times, but subsequently 4-6 hours for more sensitive detection, and viewed on a UV lightbox. The microtiter plates were divided into quadrants like the pots, and when a positive well was detected, the plants in the appropriate quadrant were flagged with numbered toothpicks and another leaf piece was tested. Once a putative mutant or mutants had been had been identified, all the other plants in that pot were removed and the putative mutant(s) was allowed to set seed. The identification system used lists the number of the pool, followed by a dash and the number of times seed had been planted from the pool, the quadrant in which the putative mutant had been found, the number of the putative mutant and finally "w" for the warm screened plants or "c" for those screened in the cold. As more work was done with a given mutant line, the designation was sometimes shortened for ease of use.

Plant growth and stress treatment

Arabidopsis thaliana (L.) Heyn. ecotype RLD or Columbia was grown in controlled environment chambers at 22°C with a 24 hour photoperiod (about 120 $\mu mol\ m^{-2}\ s^{-1}$) as previously described (Gilmour et al., 1988). Humidity was not controlled. After about three weeks of growth, plants were either harvested or subjected to one of several stress treatments. Plants were cold-acclimated by placing pots in a 2°C cold room under constant cool-white fluorescent light (45-

55 μmol $m^{-2}s^{-1}$) for four days. Drought stress was administered by excising the shoots from the pots, placing them on 15 cm Whatman No. 1 filter paper in 150 mm x 15 mm plastic petri plates with the lids open for 45 minutes. The lids were then placed on the plates and plants were incubated for four to six hours under fluorescent lights (about 50 μmol $m^{-2}s^{-1}$). Water loss was calculated by dividing the weight of the drought-stressed plants by their initial fresh weight and multiplying by 100. Plants lost from 18% to 23% of their initial weight. The ABA treatment consisted of placing excised plants into a 150 mm x 15 mm petri plate with 75 ml reverse osmosis water supplemented with 100 μM ABA (mixed isomers, Sigma) and one drop (10-20 μl) Tween-20 (polyoxyethylene-sorbitan monolaurate 20, Sigma), covering them with a piece of 15 cm Whatman No. 1 filter paper to keep them submerged, placing the lid on the plate and incubating them for 4 to 6 hours under fluorescent lights (about 50 μmol $m^{-2}s^{-1}$). Water-soaked plants underwent this same treatment, except that the ABA was omitted.

RNA extraction and fractionation

Total RNA was extracted essentially as described (Gilmour et al., 1988) with a few modifications. Frozen pulverized tissue was extracted with equal volumes of phenol/chloroform/isopropyl alcohol (25/24/1, v/v/v) and extraction buffer (1% (w/v) triisopropylnaphthelene sulfonic acid, 6% (w/v) p-aminosalicylic acid, 100 μM Tris-HCl pH 7.6, 50 mM EGTA, 125 mM NaCl, 1%SDS, 10 mM

DTT) on ice. This was further homogenized in the tube using a Tekmar® Tissumizer, centrifuged (10,000 rpm in an SA600 rotor), and the supernatant extracted again with phenol/chloroform/isopropyl alcohol. This second set of tubes and all subsequent tubes and solutions except ethanol were made free of RNase by treatment with DEPC (Sambrook et al., 1989). Nucleic acids were precipitated with cold 95% ethanol, resuspended in 1 ml double-distilled water, transferred to a microfuge tube and precipitated on ice for an hour with 1/4 volume 10 M LiCl. The pellets were precipitated again with 95% EtOH and resuspended in 200 µl water. The OD_{260} was measured to estimate the concentration of RNA.

For fractionation, RNA (5 to 40 µg) was dried down, resuspended in formaldehyde loading buffer containing EtBr (about 100 ng/ml), incubated at 68°C for 10-15 minutes and fractionated on denaturing formaldehyde agarose gels (Sambrook et al., 1989). The RNA was then transferred (Sambrook et al., 1989) to Magna NT membranes (MSI) using 10x SSPE made according to instructions provided by Schleicher and Schuell. This recipe for SSPE is slightly different from that described by Sambrook et al. (1989), and is what was used as SSPE in all experiments. RNA was UV cross-linked (Stratalinker, Stratagene) to the filters.

Northern hybridizations

For standard northern hybridizations, blots were prehybridized, hybridized and washed in a Robbins Scientific Model 400 hybridization oven according to standard methods (Ausubel et al., 1987) with some modifications. Blots were incubated with prehybridization solution, which is made up of 50% formamide, 5x SSPE, 50 mM potassium phosphate pH 8, 5x Denhardt's solution (1x is 0.02% w/v ficoll, 0.02% w/v polyvinylpyrrolidone, 0.02% w/v bovine serum albumin), 0.5% SDS, 100 μg/ml sheared, denatured fish sperm DNA, at 42°C for three hours to overnight and in hybridization solution (50% formamide, 5x SSPE, 50 mM potassium phosphate pH 6.5, 1x Denhardt's solution, 0.5% SDS, 5% dextran sulfate, 100 μg/ml sheared denatured fish sperm DNA) at 42°C overnight. Blots were rinsed and then given two 30-minute washes at room temperature in 2x SSPE/0.5% SDS, then washed two to three times, for 15 minutes each time, with 0.1x SSPE/0.5% SDS at 50°C.

Probes were made from gel-isolated fragments labeled with ^{32}P by random priming (Feinberg and Vogelstein, 1983). An *Eco* RI fragment from pHH67 was used to visualize *COR*15 message and a *Sac* I/*Hind* III fragment from pBI101.3 (Jefferson, 1987) served as a GUS probe.

Plant genomic DNA extraction and Southern analysis

Genomic DNA extraction for the analysis of the transgenes in 1,6 and for analysis of putative mutants was done as described by Rogers and Bendich (1988) except that the mixture of plant tissue and extraction buffer was incubated at 68°C for 30 minutes before chloroform extraction and the CTAB (hexadecyltrimethylammonium bromide) precipitation was left in the refrigerator until flocculant precipitate was seen.

Total *Arabidopsis* DNA or plasmid DNA was digested with various restriction enzymes and fractionated by agarose gel electrophoresis (Sambrook et al., 1989) using TAE (40 mM Tris-acetate, 1 mM EDTA). The DNA was then transferred (Sambrook et al., 1989) to Magna NT membranes (MSI) using 10x SSPE. The DNA was UV cross-linked (Stratalinker, Stratagene) to the filters.

Southern blots were prehybridized and hybridized in a Robbins Scientific Model 400 hybridization oven using the nonfat dry milk method of Johnson et al., (1984) and washed using standard conditions (Sambrook et al., 1989), except that all washes were done with 0.1% SDS/0.1x SSC. Gel-isolated fragments labeled with ^{32}P by random priming (Feinberg and Vogelstein, 1983) were used as probes. A *Sac* I/*Hind* III fragment of pBI101.3 was used to detect GUS inserts, a *Bam* HI/*Pst* I fragment of pCIB710 (Rothstein et al., 1987) was used as a probe

for the CaMV 35S promoter and an *Eco* RI fragment of pHH67 was used to visualize *COR*15a and *COR*15b sequences.

The binomial distribution $N!/n_1!n_2!\ (p)^{n_1}\ (q)^{n_2}=x$ was used to determine how many samples of putative mutant line 4-11#3w should be tested to give a 95% probability that there is no segregation from the locus thought to contain the linked inserts. N is the total number of samples, n_1 and n_2 are the number of times co-segregation and the lack of co-segregation are seen, respectively, p is the proportion of the time the lack of a band is expected from a segregating parent (1/4) and q is the proportion of the time a band should be seen from a segregating parent (3/4). The percent chance of not seeing segregation when there really was some is x. The exclamation mark denotes the factorial of the number preceeding it. The equation was solved for n_2 after n_1 had been set to zero and x had been set to 5% or less. This yielded an n_2 of eleven. Since n_1 is zero, that means N is also eleven and eleven samples must be tested.

Literature Cited

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current Protocols in Plant Molecular Biology. Greene Publishing Associates and Wiley Interscience, New York, Chichester, Brisbane, Toronto, Singapore

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. 24: 701-713

Bouchez D, Camilleri C, Cabouche M (1993) A binary vector based on Basta resistance for *in planta* transformation of *Arabidopsis thaliana*. C. R. Acad. Sci. 316:1188-1193

Bowling SA, Guo A, Cao H, Gordon AS, Klessing DF, Dong X (1994) A mutation in *Arabidopsis* that leads to constitutive expression of systemic acquired resistance. The Plant Cell 6:1845-1857

Brusslan JA, Karlin-Neumann GA, Huang L, Tobin E (1993) An *Arabidopsis* mutant with a reduced level of *cab140* RNA is a result of cosuppression. The Plant Cell 5:667-677

Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. The Plant Cell 6:1583-1592

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6-13

Gilmour SJ, Hajela RK, Thomashow MF (1988) Cold acclimation in *Arabidopsis thaliana*. Plant Physiol. 87:745-750

Hajela RK, Horvath DP, Gilmour SJ, Thomashow MF (1990) Molecular cloning and expression of *cor* (Cold-regulated) genes in *Arabidopsis thaliana*. Plant Pysiol. 93:1246-1252

Haughn GW, Somerville CR (1986) Sulfonurea-resistant mutants of *Arabidopsis thaliana*. Mol. Gen. Genet. 204:430-434

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol Rep. 5:387-405

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. The EMBO Journal 6:3901-3907

Klee HJ, Horsch RB, Hinchee MA, Hein MB, Hoffman NL (1987) The effects of overproduction of two *Agrobacterium tumefaciens* T-DNA auxin biosynthetic gene products in transgenic petunia plants. Genes Dev. 1:86-96

Levitt J (1980) Responses of plants to environmental stress. Chilling, freezing, and high temperature stresses, second edition. Academic Press, New York

Moffat B, Somerville C (1988) Positive selection for male-sterile mutants of *Arabidopsis* lacking adenine phosphoribosyl transferase activity. Plant Physiol. 86:1150-1154

Negrutiu I, Beeftink F, Jacobs M (1975) *Arabidopsis thaliana* as a model system in somatic cell genetics. Plant Sci. Lett. 5:293-304

Nussaume L, Vincentz M, Caboche M (1991) Constitutive nitrate reductase: a dominant conditional marker for plant genetics. The Plant Journ. 1:267-274

Perera RJ, Linard CG, Signer ER (1993) Cytosine deaminase as a negative selectable marker for *Arabidopsis*. Plant Mol. Biol. 23:793-799

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues. In SB Gelvin, RA Schilperoort, eds., Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, Boston, London, pp A6/1-11

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A laboratory manual, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Susek RE, Ausubel FM, Chory J (1993) Signal transduction mutants of *Arabidopsis* uncouple nuclear *CAB* and *RBCS* gene expression from chloroplast development. Cell 74:787-799

Takahashi T, Naito S, Komeda Y (1992) The *Arabidopsis* HSP18.2 promoter/GUS gene fusion in transgenic *Arabidopsis* plants: a powerful tool for the isolation of regulatory mutants of the heat-shock response. The Plant Journal 2:751-761

Thomashow MF (1993) Genes induced during cold acclimation in higher plants. In, PL Steponkus, ed., Advances in low temperature biology, vol. 2. JAI Press Ltd., London, pp 183-210

Thomashow MF (1994) *Arabidopsis* as a model for studying mechanisms of plant cold tolerance. In, EM Meyerowitz, CR Somerville, eds., *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 807-834

Valvekens D, Van Montagu M, Van Lijsebettens M (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. PNAS 85:5536-5540

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119

Chapter 3:

Analysis of a genomic clone containing *COR*15a and *COR*15b

Summary

In order to facilitate the interpretation of studies of the regulation and protein analysis of the cold-regulated gene *COR*15, experiments were carried out to determine whether it was a member of a small gene family. Initial genomic Southerns of *Arabidopsis* had shown that *COR*15 was either a single gene or a member of a small gene family (Hajela RK, Horvath DP, Gilmour SJ, Thomashow MF, 1990, Plant Physiol. 93:1246-1252) and further genomic Southern analysis indicated that there was a second gene which hybridized to a cDNA corresponding to *COR*15. A genomic clone containing *COR*15 was found to also contain this second member of the *COR*15 gene family, which was then named *COR*15b. *COR*15 was renamed *COR*15a. The two genes are in the same relative 5′ to 3′ orientation and their predicted coding regions are 82% identical at the nucleic acid sequence level. COR15a and the protein predicted to be encoded by COR15b were compared to each other and a model for their

secondary and tertiary structure is proposed. *COR*15b is inducible by low temperature and possibly by ABA, but not by drought. Gene fusion experiments indicate that this lack of drought inducibility is most likely promoter-based.

Introduction

Arabidopsis thaliana is able to increase in freezing tolerance at low, nonfreezing temperatures in a process known as cold acclimation (Gilmour et al., 1988). During this process, numerous genes are induced, including a gene known at the beginning of this work as *COR*15 (Hajela et al., 1990). *COR*15 mRNA is inducible by drought stress and by the exogenous application of abscisic acid as well as by low temperature, but not by heat shock. In vitro translation experiments showed that *COR*15 encodes a 15 kDa polypeptide, COR15, which is cleaved to 9.4 kDa upon import into chloroplasts. The mature form of COR15, COR15m, is acidic, with a pI of 4.6, and is very hydrophilic. COR15m remains soluble upon boiling, a property thought to be a consequence of this high hydrophilicity (Lin and Thomashow, 1992).

A genomic clone containing *COR*15 was isolated in preparation for gene regulation studies (see Chapter 2). Approximately 3 kb of the *COR*15 promoter was identified in this clone and subsequent promoter deletion analysis showed

that the cold- drought- and ABA- inducibility of *COR*15 is promoter-based (Baker et al., 1994).

In the analysis of *COR*15, the question of whether it was a single gene or a member of a small gene family was raised. *COR*6.6 and *COR*78 had been found to be members of small gene families (Kurkela and Borg-Frank, 1992 and Yamaguchi-Shinozaki and Shinozaki, 1993) and an extra band found on westerns when examining the translated product of *COR*15 might be explained by the presence of two related genes. Initial analysis using two restriction enzymes gave a simple pattern, suggesting that COR15 was either a single gene or was a member of a small gene family (Hajela et al., 1990). The restriction map of the genomic clone isolated allowed for a more efficient detailed analysis, so more genomic Southerns were performed, and evidence was found for the presence of a second gene.

The genomic clone already isolated containing *COR*15 was found to also contain this second gene, so *COR*15 was renamed *COR*15a and the second gene was designated *COR*15b. This chapter describes the sequencing of 4.2 kb of this clone, analysis of *COR*15b sequence, a knowledge-based model of the protein it encodes, and analysis of *COR*15b inducibility.

Results

Identification of *COR*15b and sequence comparison with the coding region of *COR*15a

Southern analysis of *Arabidopsis* (ecotypes Columbia and RLD) genomic DNA was performed under conditions of high and low stringency to look for the presence of a second gene. A cDNA clone of *COR*15, pHH67 (Hajela, 1990), was used as a probe. Four restriction enzymes (*Sca*I, *Hin*dIII, *Nco*I and *Eco*RI) were chosen to give a predictable banding pattern, based on a restriction map of the genomic clone. Indeed, in both ecotypes faint bands were seen in addition to the expected bands on the low stringency blot for three of the four enzymes used (data not shown). The absence of extra bands with the fourth enzyme, *Eco*RI, which should give the largest fragment containing *COR*15, suggested that both genes might be contained within this fragment, and thus both be present on the genomic clone already isolated (see Chapter 2). Closer examination of preliminary sequence about one kilobase downstream of *COR*15 did reveal identity with the promoter of *COR*15. This second gene was designated *COR*15b, and the original gene renamed *COR*15a.

Further sequence analysis was carried out. The sequence of approximately 4.2 kilobases, beginning about one kilobase upstream of the start of transcription for *COR*15a and extending 448 bases past the stop codon of *COR*15b is shown (Figure 3.1). The region of the genomic clone that corresponded to cDNAs of

agatcttttc cgttgaattt attttagact ttttttttaa tggacttcat tttaaatttt 60
tacaaaatta aattattgca ttttctattc atattgaatt aggagatgtt actgtccgtc 120
agattctcta gacttttttt tttaaagact gatctatgat cagaattcca attttttttt 180
tctttaagga aatacatcag agagaaaaat tattacgaaa cgattctatt acaagtaatg 240
attttaacct ttttttttt acaattgaca atcttttcac aacaaaaatc cacaagaaac 300
gttagacaat ggcataaatt tatttaaatt aatccgtata tattcgcctt ctatgagaat 360
tgaattctat accactgtaa aattcttaaa cgagataaga ttattttcag catgtaaaaa 420
atggtttgtg gtttcaactc atttgggcta ttagttttac atttaggctt gcaaccttgt 480
cggtttattt tgtgtaggct tttggtagat ttgggcttgc aaacccaaat taacttgttg 540
gccgacatac atttgtttct attacaaatt taacaacaaa cgtcaataaa tacacgtgaa 600
ggaaatgaga acgaccctct taagtagtac tggaaattga aaaaaagaaa tctagaaatg 660
ctaacatgta agtttttgtt accaaaaatg caatttgtat gtagccacaa tttcatggcc 720
gacctgcttt ttttttcttc ttctttctga aaaccacaaa tatgattaca cgtggcctga 780
aaagaacgaa cagaaactcg gtaatgtgca aaaaatatct tactcttaat acgtgtaatt 840
ttggagtgta ataggtctat cgatctataa aacgatacta ttggagatta gattcttctc 900
P+1 *COR15a*
atctcacttt gttcatctaa aaactcctcc tttcatttcc aaacaaaaac ttctttttat 960

tctcacatct taaagatctc tctcATGGCG ATGTCTTTCT CAGGAGCTGT TCTCACTGGT 1020

ATGGCTTCTT CTTTCCACAG CGGAGCCAAG CAGAGCAGCT TCGGCGCTGT CAGAGTCGGC 1080

CAGAAAACTC AGTTCGTCGT CGTTTCTCAA CGCAAGAAGT CGTTGATCTA CGCCGCTAAA 1140

GGTGACGGCA ACATCCTCGA TGACCTCAAC GAGGCCACgt aagtctacat tctttcttct 1200

tttagtatct tgcctcataa gtaaggatct tagcaggcaa tgtttatggt atactatatt 1260

agtatagatt ttagtggaaa tatgtttgtt ttgaacttat tttatgatca tatttgacta 1320

ttatcaaaga taaagattca tataccgtac attatatatc tctatttttc tagtttacat 1380

gtatagctct aagtttattt gatgattctg ttgactactt ttggatatgt gttttgaaac 1440

ctttgataaa tactaaaata aattttaatt tgaaaatgat atagAAAGAA AGCTTCAGAT 1500

TTCGTGACGG ATAAAACAAA AGAGGCATTA GCAGATGGTG AGAAAGCGAA AGACTACGTT 1560

GTTGAAAAAA ACAGTGAAAC CGCAGATACA TTGGGTAAAG AAGCTGAGAA AGCTGCGGCG 1620

TATGTGGAGG AGAAAGGAAA AGAAGCCGCA AACAAGCCGG CAGAGTTCGC CGAGGGTAAA 1680

GCAGGAGAGG CTAAGGATGC CACAAAGTAG ggtcttacct aatcagttaa tttcaagcac 1740

ttaaactcgt agatatattg atccatatcc tctctcttca tgtttaatag tacttacaat 1800
aagatgagtc tgttgtaatt tctattaatt tcacatcaca actgaaataa gatatggtat 1860
ccacagtcac cgtcacattc tttatgtttt gcaaaatatt caatagacaa attaaataat 1920
gagtaattat aagtaacatc gccgtctaga gtctttcaac ggtgaaagga tgagagcata 1980
tgggagatcg atagccttgt ctcttttgtg acgaaccaat aaaagaggaa ctgctttttt 2040
tcttcgcatg tccactattt actgtttggt atgacatcga attggtatga tacattgttg 2100
taccaaaaac ttcaaaattt ggtggcttgt aatcaacatt gatccactta gccattccaa 2160
gcagtgaatt acaatccgtg gaaggaaaaa ttatggtccg gcaccaacaa attatcacct 2220
catctcccgt cacaaacata tcatctagaa gtcggtatgc cacaccacat aattaaggag 2280
cacattactc acataccaa ccaatgtggg acatatctaa taagctcatt cttggctggt 2340
acatgctcta tttcatattt atcaaaaaaa aaaatattag gcattgtaaa tagcgttttt 2400
gctgttgagc aaaatagtta tatttgaaag taacattggt atttataatt atagtataac 2460
aattaggcat tgaagtgtga gtttttgttt ttgtttattt aacattggag tattaggttc 2520
ttagaaatat atctatatac tattagtagt ttaactacag tttgtactta attgaaaaaa 2580

```
tgttaaaagt tgttttaacc tagctaattg ctaaaaatga ctaaatagac atacacaaag 2640
acttgtacat tttcagctta acgactaata cattttttcct ttatatatat atctctatcg 2700
agtctagtta ttaatgttga aagttgcaaa taaaacagaa atgctaacat gtaaatatcg 2760
tagccaaaaa tgctaacatg tgtataacgg ttataaccac aacttgatgg ccgacctctt 2820
ttttcttttg gtaaccatag aaatggttac acgtaactag tacgaaccaa cgaaaactct 2880
tcttattcga tagttaaaga taatagcaat gcgcaaaaat atctagcact cacacgtgta 2940
gttttggatt ctcattggtc gagagatcta taaaacgata ctattggagg ttagattttt 3000
   +1 COR15b
ctcatctcac tttctccatc ttaaaactct ttcttgtatt tattttcctc ccaaaaaaca 3060

tctttaagag tcctcATGGC GATGTCTTTA TCAGGAGCTG TTCTCAGTGG CATGGGTTCT 3120

TCTTTCCACA ACGTAGGAGC AAAGCAGAGT GGTGTTGGTA CCGTCAGAGT TGGCCGGAAG 3180

AGTGAGCTCG TCGTCGTTGC TCAGCGCAAG AAGTCGTTGA TATACGCCGT TAAAAGTGAC 3240

GGCAACATCC TCGATGACCT CAACGAAGCC ACgtgagtct atattctttt cttttaacaa 3300

attttctcgt gactaaggct atcttcatgt gtactcgcaa aaactctcat atacattacg 3360

tactcgccat gatagttatt ttattttggt taatcccatc ataacagttt tatacatact 3420

ataatcaact ataatggcct tgttgaggat tgagtttcca tggtgtagat atatgttttt 3480

aaaatcgttt gtcaataatg aatcttaaca tatatatata tatatatagA AAGAAAGCTT 3540

CTGATTTCGT GACGGATAAG ACGAAGGAGG CCTTGGCGGA TGGCGAGAAA ACAAAAGACT 3600

ACATTGTTGA GAAAACCATT GAAGCCAATG AAACTGCGAC TGAGGAAGCT AAGAAAGCTT 3660

TGGATTATGT GACTGAGAAA GGAAAAGAAG CCGGAAACAA GGCGGCTGAG TTCGTAGAGG 3720

GTAAAGCAGA AGAGGCTAAG AATGCCACAA AGTCCTGAtt caaccactca tgtagtcatg 3780

ttcaatagtt tacaataaga tcagtttgtt gtaatttctt tttcaccatt tgaaatgaag 3840
agctgttaat tactcactag cacattttaa atacgtatca aatgatttaa atataaacta 3900
cattttagtg actacaagta acatagtagt atttatgagt tggattcggg aaaaatctgt 3960
caaaatcttc tctcattttg atcctcaatt ataaagaaac tgtgaaactg atgactgttt 4020
ttatgacaac ccctaataat agcctaaaaa gatcaaaaca cctaaaacat caacaattgt 4080
tatatatcca tctcctgagt gaaatctcat aggtgtaagg ggcaatacat actctctttg 4140
tgaatgtaga agaaatattt tttttgtcaa aagttgtaac caaaacaaaa agaacattgg 4200
aagctt                                                          4206
```

Figure 3.1 Sequence of a 4.2 kb region of genomic DNA containing *COR*15a and *COR*15b. Noncoding sequences, including promoter regions, introns and 3′ nontranslated regions are shown in small letters, while all translated sequences are in uppercase. The arrows show the start of transcription for each gene.

*COR*15a was only sequenced on one strand, since it exactly matched the sequence of the cDNA clones. This confirmed that the intron found in one of the cDNA clones (pHH71.1) is the only one present in *COR*15a and that there are no other differences between the cDNA and genomic versions of *COR*15a. The rest of the sequencing was done on both strands. The two genes are separated by less than one kilobase of DNA and are in the same relative 5′ to 3′ direction (Figure 3.1). The transcriptional start site of *COR*15a was determined by primer extension (Baker et al., 1994), while that of *COR*15b is predicted based on comparison to that of *COR*15a and on the consensus sequence found at transcriptional start sites (Joshi, 1987). The location of the intron of *COR15a* is known based on comparison to the sequence of the cDNA clone pLCT10A (Lin and Thomashow, 1992), while the intron of *COR*15b is predicted based on sequence comparison to *COR*15a and on consensus sequences found at intron/exon boundaries (Brown, 1986). The predicted coding regions and introns are, respectively, 82% and 57% identical at the nucleic acid level.

Analysis of the protein predicted to be encoded by *COR*15b and comparison with COR15a

The polypeptide predicted to be encoded by *COR*15b, designated COR15b, is 77% identical and 82% similar in amino acid sequence to COR15a and both have four degenerate 13 amino acid repeats (Figure 3.2). Like COR15a, which is known to be targeted to chloroplasts (Lin and Thomashow, 1992), the N-terminal sequence of COR15b has features typical of chloroplast transit peptides (Gavel

```
COR15b  1 MAMSLSGAVLSGMGSSFHNVGAKQSGVGTVRVGRKSELVVVAQRKKSLIYA*V 52
COR15a  1 ....f.....T..a....-s.....sf.a....q.Tqf...s......,..*A 51

       53 KSDGNILDDLNEATKKASDFVTDKTKEALADGEKTKDYIVEKTIEANETATEE 105
       52 .g.....,.....,.....a...V...ns.taD.Lgk. 104

      106 AKKALDYVTEKGKEAGNKAAEFVEGKAEEAKNATKS                  141
      105 .e..Aa..e......a......A....g...d...                   139
```

Figure 3.2 Alignment of COR15a and COR15b. Identical, biochemically similar and different amino acids are indicated by dots, capital and lower-case letters, respectively. The consensus cleavage site sequence for chloroplast transit peptides (Joshi, 1987) is double-underlined, with the predicted site of cleavage being indicated by an asterisk. The four copies of the 13 amino acid repeat are underlined. One gap, indicated by a dash, was introduced to optimize the alignment.

and von Heijne, 1990). It has a high serine/threonine content (18%), only one acidic residue (Glu), and a sequence beginning at amino acid 49, IYAV, that resembles the loosely defined chloroplast signal sequence cleavage site V/I-X-A/C-A (Gavel and von Heijne, 1990). While the final valine of COR15b's predicted cleavage site is absent from the consensus, it is found in a number of the sequences from which the consensus was derived. Finally, there is an arginine residue eight amino acids before this putative cleavage site, as would be expected. It is, however, also possible that COR15b may be targeted to mitochondria, since mitochondrial transit peptides have much in common with chloroplast transit peptides (Soll and Alefsen, 1993). The predicted molecular weight of the mature COR15b protein, if it is targeted and processed, is 9.6 kDa, compared to 9.4 kDa for the mature form of COR15a (Lin and Thomashow, 1992). Both are quite acidic, each with a predicted pI of 4.6.

Proteins similar to COR15

The BLAST algorithm tblastn (Altschul et al., 1990; blast@ncbi.nlm.nih.gov) was used to search nonredundant databases, including GenBank (Release 92, December 15, 1995) and the EMBL data library (Release 45.0), for sequences similar to the predicted mature COR15b polypeptide (COR15bm). The putative mature forms of BN115, BN19 and BN26 from *Brassica napus* (accession numbers U14665, S68879 and S68727, respectively; Weretilnyk et al. 1993) and of a *Brassica oleracea* cold responsive mRNA (accession number U16751; Jo H, Lee C, and

Sohn U, unpublished data) show amino acid sequence identity of 72.2, 68.9, 66.7 and 70%, respectively, with COR15b. The other genes with the highest sequence identity identified by this search are Group 3 late embryogenesis abundant (LEA) proteins isolated from birch (accession number Z18891; Puupponen-Pimia et al., 1993), *Arabidopsis* (accession number D64140; Yang H, Saitou T, Harada H and Kamada H, unpublished data), soybean (accession number U02966; Chow T, Hsing YC, Chen Z, unpublished data, and accession number Z22872; Hsing et al., 1996, In press), carrot (accession number X16131; Franz et al., 1989) and wheat (accession number M72395; Curry and Walker-Simmons, 1993). These have 20 to 40% amino acid sequence identity to *COR*15b, primarily in repeat regions characteristic of Group 3 LEA's (Dure, 1993), which align with the degenerate repeats of COR15a and COR15b. Neither the structure nor the function is known for any of these proteins

Modelling of COR15 structure

COR15bm also shows similarity to COR15am when structural modelling is done. The secondary structures of both COR15am and COR15bm are predicted by PHDsec (Figure 3.3A) to be primarily alpha-helical with short loops at each terminus and two potential turns or unstructured regions separating helices (Rost and Sander, 1993 and Rost and Sander, 1994). This method aligns the sequence of interest with similar sequences from the SWISSPROT database and uses this alignment to develop a prediction of secondary structure using the

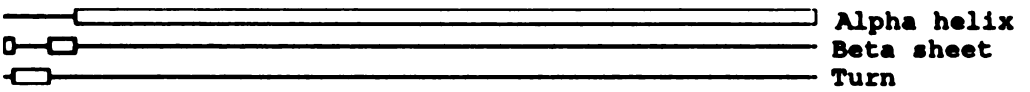
A

```
 AA VKSDGNILDDLNEATKKASDFVTDKTKEALADGEKTKDYIVEKTIEANETATEEAKKALD
sec LLL.HHHHHHHHHHHHHHHHHHHHHHHHHH.......HHHHHHHHHHHHHHHHHHHHH
prob 9992799999999889899999998999974441223899999988899999999999

 AA YVTEKGKEAGNKAAEFVEGKAEEAKNATKS
sec HHH...HHHHHHHHHHHHHHHHHH..LLL
prob 997434768999999999986687533889
```

B

```
 AA VKSDGNILDDLNEATKKASDFVTDKTKEALADGEKTKDYIVEKTIEANETATEEAKKALD
sec LLLLLLL..HHHHHHHHH...HHHHHHHHHH.............L....HHHHHHHHH
prob 99999864379999987722157999999991224121221223254131899999999

 AA YVTEKGKEAGNKAAEFVEGKAEEAKNATKS
sec HHHHHH....HHHHHHHH..HHHHH..LLL
prob 99999813426899999715589822699
```

C



D

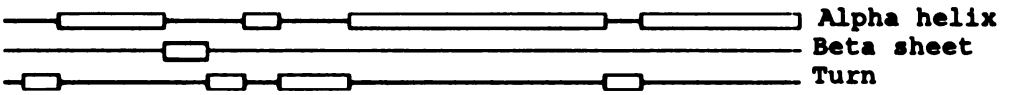


Figure 3.3 Predictions of the secondary structure of COR15bm. A. PHDsec secondary structure prediction based on alignment of COR15bm with proteins found in a search of the SWISSPROT database. B. PHDsec secondary structure prediction based on alignment of COR15bm with COR15am and the predicted mature forms of BN115, BN19 and BN26. For the PHDsec predictions, amino acid sequence (AA), secondary structure prediction (sec) and probability (prob) are shown. PHD predicts loops (L), helices (H) and beta sheets (E, for extended), with 1 being the lowest and 9 the highest probability of accuracy. Dots show where the prediction is uncertain. Lines over PHDsec predictions denote helices predicted by knowledge-based modelling using beta-lactamase from *Bacillus licheniformis*. C. Secondary structure prediction based on the algorithm of Garnier and Robson. D. Secondary structure prediction based on the algorithm of Chou and Fasman. For the Garnier-Robson prediction and Chou-Fasman predictions, open boxes indicate the presence of the structural feature and lines indicate its absence.

profile neural network method (Rost and Sander, 1993). A neural network is essentially a system of artificial intelligence which can be trained in pattern recognition (Rost and Sander, 1993). The variables which are optimized during training are internal to the program, so the criteria on which the prediction are based are not available for examination. However, it is known that the amount of sequence conservation in the aligned sequences aids PHDsec in predicting secondary structure. This prediction method is rated at an expected 72.1% average accuracy for the three states of helix, strand and loop (Rost and Sander, 1993). When the same program used proteins known to be highly similar to COR15bm in its alignment, namely COR15am and the predicted mature forms of BN115, BN26 and BN19, the prediction was still highly helical with loops at both ends, but had more regions of uncertain structure (Figure 3.3B).

Prior modelling using the less accurate algorithms of Garnier and Robson (Figure 3.3C) (Garnier et al., 1978) and Chou and Fasman (1978) (Figure 3.3D) had predicted the high helicity, but only the method of Chou and Fasman had predicted turns separating helices. It was not expected that COR15m would form a single long helix, since this is unusual for a soluble protein and has only been observed in calcium-binding and fiber-forming proteins (observation from protein structural analysis by LA Kuhn and JA Tainer). When COR15bm is modelled as a single helix in helical wheel format (Schiffer and Edmundson, 1967), it is notably amphipathic (Figure 3.4). One side of the helix is notably enriched in hydrophobic amino acids, while the other is highly hydrophilic.

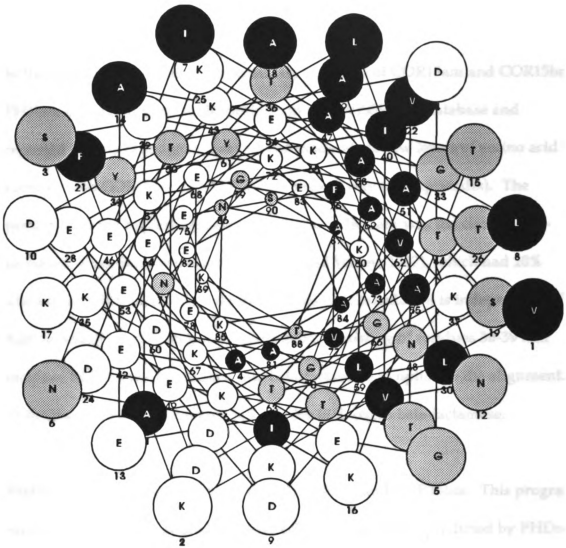


Figure 3.4 COR15bm modeled as an amphipathic helix. Single letter codes of the amino acids are shown. Hydrophobic residues (A, L, I, V and F) are shown as filled circles, polar residues (S, T, G, N and Y) are shaded and charged residues (D, E and K) are white.

Helical net modelling (Dunnill, 1968) further reveals potential ion pairs (Figure 3.5) that may be involved in stabilizing the helix or helices (Marqusee and Baldwin, 1987).

In the process of predicting the secondary structure of COR15am and COR15bm, PHDsec compared them to the proteins in the SWISSPROT database and reported 21 and 12 proteins that showed moderate to low levels of amino acid identity with COR15am and COR15bm, respectively (see Appendix). The protein with the highest sequence identity to COR15am was the class A beta-lactamase from *Bacillus cereus* (Madgwick and Waley, 1987), which had 28% identity and 38% similarity over 89 residues (this calculation is independent of that provided by PHDsec). Two short gaps, consisting of residues 36-39 and residues 64-70, were introduced into beta-lactamase to optimize the alignment. COR15bm is 23% identical and 34% similar to *B. cereus* beta-lactamase.

PHDthreader was also used to analyze COR15am and COR15bm. This program combines the secondary structure and solvent accessibility predicted by PHDsec and PHDacc (a solvent accessibility prediction program which also uses a neural network) with amino acid identity of a given protein into a string of information that utilizes 120 characters (20 amino acids multiplied by three secondary structure states and two states of solvent accessibility). This string of information is then compared with the corresponding string of amino acid

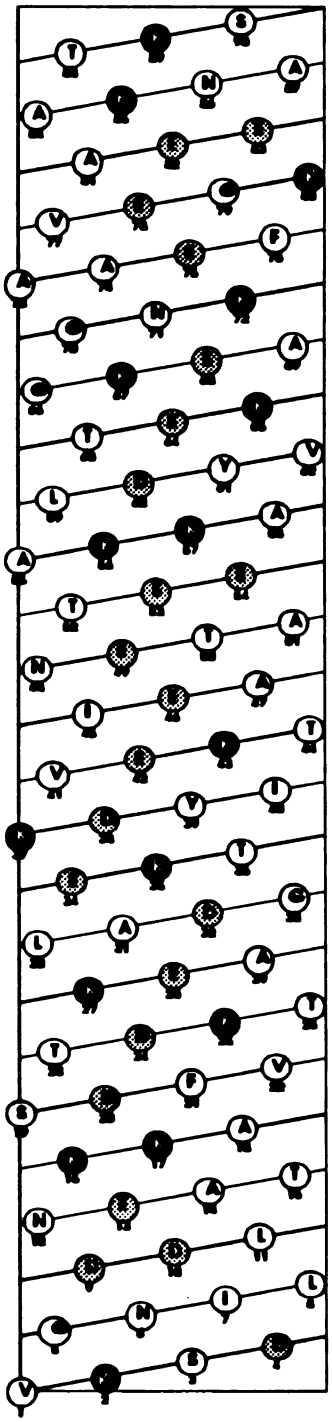


Figure 3.5 Helical net model of COR15bm. Single letter codes of the amino acids are used. Positively charged residues are shown as black circles and negatively charges residues are shaded. Hydrophobic and polar residues are white. The proximity to each other of positively and negatively charged residues on successive turns of the helix may allow intrahelical electrostatic interactions. This could stabilize the helical conformation.

identity, secondary structure and solvent accessibility calculated from each of 450 nonredundant proteins of known tertiary structure in order to look for potential remote homologs (proteins with similar tertiary structure, but no significant sequence similarity to the protein of interest) (Rost 1995a and Rost 1995b). PHDthreader returned an alignment of 20 potential remote homologs of known structure for COR15am and for COR15bm (see Appendix). For both COR15am and COR15bm, the aligned protein with the highest sequence similarity was apolipophorin III from the African locust *Locusta migratoria* (Breiter et al., 1991). However, the sequence identity of apolipophorin III with COR15am is only 22% and with COR15bm is 16%, and there is one five-residue gap in the alignment. Two other factors discouraged use of apolipophorin III in modelling. One was the low reliability of the threader technique. The method is only rated at 30% accuracy and the information on threading given at the world wide web site that carries PHDthreader warns that predictions will likely be wrong. The second was that the location of the region of apolipophorin III that is similar to the entirety of COR15am and of COR15bm maps to three out of five helices that form a helical bundle. Apolipophorin is 67 and 66 residues longer than COR15am and COR15bm, respectively. The absence of the other two helices is likely to affect the way the protein folds.

Of the reported proteins with similarity to COR15am and COR15bm for which the three-dimensional, atomic resolution structure is known, beta-lactamase was the most similar, making it the best candidate for knowledge-based modelling.

It has been shown that proteins having greater than or equal to 25% identity over an 80 residue stretch are likely to have greater than 70% tertiary structure identity (Sander and Schneider, 1991), although the gap opening penalty of 3.0 and gap elongation penalty of 0.1 reduce the effective identities of COR15am and COR15bm with *B. cereus* beta-lactamase to 20.9% and 16.9% respectively. Furthermore, COR15am and COR15bm are expected to have essentially the same structure, based on their 77% amino acid sequence identity with no gaps or insertions. In a study of the relationship between sequence identity and structural similarity for 51 proteins (Hilbert et al., 1993), it was found that the alpha carbons of proteins with 77% sequence identity will superimpose to within 0.5 Å RMSD (root mean squared deviation). In other words, the structures of COR15am and COR15bm should overlap. This means that it should be reasonable to model COR15bm, as well as COR15am, on beta-lactamase.

Because the sequence identity between COR15am and beta-lactamase was on the border of Sander and Schneider's threshold (1991) of 25% identity over 80 amino acids, modelling was continued with caution with the understanding that the resulting structure could be evaluated for reasonableness based the following

four criteria. First, does the modelled structure correspond with the PHDsec secondary structure prediction? Second, are the majority of hydrophobic residues buried and hydrophilic residues exposed? Third, are insertions in COR15bm relative to beta-lactamase located in loop or turn regions of the surface rather than in the middle of helices? Finally, does the model comprise a compact globular subdomain of beta-lactamase rather than extended, poorly packed segments of secondary structure?

Because the three-dimensional structure of the beta-lactamase from *B. cereus* is not available, and there are several structures available for other class A beta-lactamases, the structural conservation among these class A beta-lactamases was evaluated. The three-dimensional structures of Class A beta-lactamases from *Staphylococcus aureus* (1blc, Chen and Herzberg, 1992) and *Bacillus licheniformis* (4blm, Knox and Moews, 1991), were acquired from Protein Data Bank (Abola et al., 1987). When these structures and the structure of the class A beta-lactamase from *Escherichia coli* (Strynadka et al., 1992) were superimposed using the protein modelling program InsightII (version 95.0), they were seen to have very similar structures (Figure 3.6). Some of the helices are at slightly different angles and some loop regions vary slightly, but the overall structure is the same.

In quantification of this similarity, the class A beta-lactamases from *S. aureus* and *E. coli* are 24.2% identical, but have an RMSD of the alpha carbons of only 1.6 Å.

Figure 3.6 Superimposition of three Class A beta lactamases. The structures of Class A beta lactamases isolated from *Staphylococcus aureus* (yellow ribbon), *Escherichia coli* (blue ribbon) and *Bacillus licheniformis* (pink ribbon) were superimposed using Insight II (version 95.0).

Those from *S. aureus* and *B. licheniformis* are 36.7% identical with an RMSD of 1.4 Å, and those from *E. coli* and *B. licheniformis* are 28.9% identical with an RMSD of 0.7 Å. Sander and Schneider (1991) use an RMSD of 2.5 Å as their upper limit for acceptable structural similarity, so these numbers are well within this limit. Thus, while other class A beta-lactamases may have slight differences in the angles of helices or loops, they are highly likely to have the same secondary structures and tertiary fold as the beta-lactamase from *B. cereus*. This high level of structural similarity among beta-lactamases, given the moderate sequence identity, also supports the idea that the structure of the COR15m proteins may be quite similar to that of beta-lactamase, despite the moderately low sequence identity.

Class A beta-lactamases of known structure were then evaluated for their sequence similarity with COR15am and COR15bm. The beta-lactamase from *Bacillus licheniformis* (Knox and Moews, 1991) had the highest identity with both COR15m proteins, having 15% identity and 29% similarity to COR15am and 16% identity and 28% similarity to COR15bm, so it was chosen for use in modelling COR15bm.

Using InsightII, the side chains of *B. licheniformis* which varied in the alignment with COR15bm were swapped for the corresponding COR15bm side chains. The rest of the beta-lactamase polypeptide not analogous with COR15bm was then

deleted from the COR15bm model. The resulting protein is comprised of a three helix bundle which has a stretch of relatively unstructured amino acids (Figure 3.7). The amino terminal seven amino acids of the model have been omitted, since they form a helix in beta-lactamase, but are strongly predicted to be a loop in secondary structure analyses of COR15am and COR15bm. These residues are in a region of no identity between the two proteins and, if they were modelled as an amphipathic helix, its hydrophobic residues would be exposed to the solution, which would be thermodynamically unlikely. It is more likely to be a loop region that packs its two hydrophobic residues against the hydrophobic regions of the protein.

The model was then evaluated according to the criteria previously proposed. First, the placement of helices in this model is similar, although not identical, to the PHDsec prediction which utilized a provided alignment of COR15am and the predicted mature forms of BN115, BN19 and BN26 (see bars over PHDsec predictions in Figure 3.3). The high sequence identity of the proteins used in this prediction should increase its accuracy over that of the prediction based on proteins of lower identity found in the SWISSPROT database. Second, hydrophobic residues are generally buried in the structure, while hydrophilic residues are surface-exposed. Third, residues that are inserted in COR15bm relative to beta-lactamase occur in surface-exposed loops (see Figure 3.7) and are thus not likely to have a major effect on the overall structure. Based on

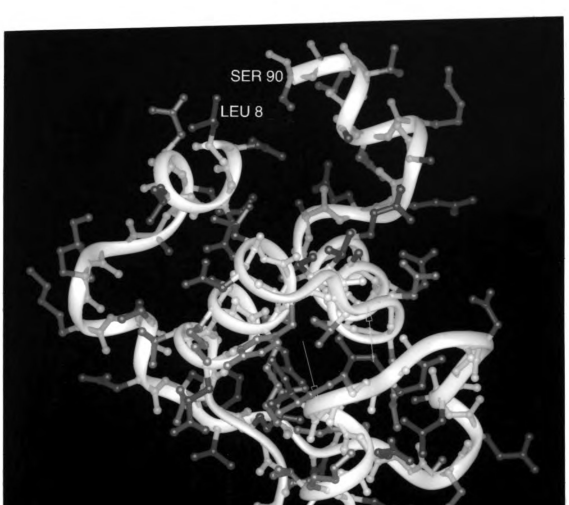


Figure 3.7 Model of the proposed structure of COR15bm. Shown is the ribbon of residues B78-B151 of the beta lactamase from *B. licheniformis*. The sidechains of the *B. licheniformis* protein have been replaced by the sidechains from COR15bm, starting with COR15bm residue 8 (Ile) and ending with residue 90 (Ser). Two yellow arrows at the center of the picture identify sites where residues of COR15bm were looped out of the alignment, and thus out of the model. The arrow on the left marks the point where residues 36-39 of COR15bm were omitted, and the arrow on the right shows where residues 64-70 were left out. Sidechains are colored according to their properties. Acidic residues (D and E) are red, basic residues (K) are blue, polar residues (S, T, G, N and Y) are magenta and hydrophobic residues (A, L, I, V and F) are green.

comparisons with the effects of insertions in other proteins, these insertions would likely extend the loop lengths or possibly extend the helix lengths (Heinz et al., 1993). Finally, it can be seen that the modelled region corresponds to a single, compact subdomain of the beta-lactamase protein (Figure 3.8).

It is likely, however, that the sidechains are somewhat more tightly packed in the actual protein, since the modelled COR15bm has a surface loop formed by the mainchain in which the sidechains are less tightly packed in the modelled COR15bm than in beta-lactamase. This results in there being a tunnel through the periphery of the modelled COR15bm, while there is no tunnel through the corresponding loop of beta-lactamase.

Regulation of *COR*15a and *COR*15b mRNA

The finding that there are two very similar members of the *COR*15 family led to the question of why the duplication might be beneficial. Are the two genes redundant or might one be present at a different time or place than the other? In order to examine the inducibility of *COR*15a and *COR*15b individually, oligonucleotides complementary to a divergent region of the 3′ end of *COR*15a and *COR*15b were synthesized. Once the hybridization conditions at which these probes were specific had been empirically determined, they were used on northern blots prepared from plants exposed to low temperature, drought or ABA (Figure 3.9A). Both genes were induced by low temperature and treatment

Figure 3.8 Region of beta lactamase with similarity to COR15am and COR15bm. Shown is a ribbon depiction of beta lactamase from *B. licheniformis* with the region that has amino acid sequence similarity with COR15am and COR15bm colored yellow.

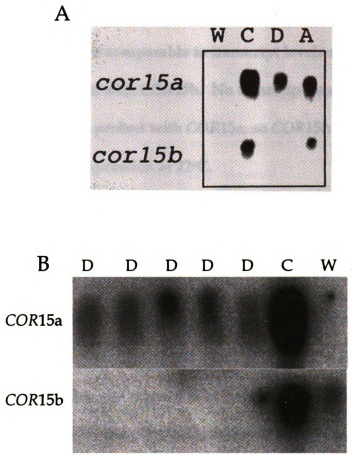


Figure 3.9 Comparison of the regulation of *COR*15a and *COR*15b. A. RNA from warm- (W), cold- (C), drought- (D) and ABA- (A) treated plants was fractionated on duplicate gels and blotted. Each blot was probed with a gene specific probe for *COR*15a or *COR*15b. B. Five drought-treated (D) RNA samples with cold (C) and warm (W) controls. Again, duplicate blots were prepared and each was hybridized with a probe specific for *COR*15a or *COR*15b.

with ABA. However, only *COR*15a was notably responsive to drought. This result was confirmed with five more drought samples (Figure 3.9B). Induction was observed when they were probed with *COR*15a, but not with *COR*15b. There were low levels of *COR*15b transcript detectable in one of the drought samples, but these were comparable to transcript levels seen in some warm control samples probed with *COR*15b. No transcript was detected in these same warm control samples probed with *COR*15a, so *COR*15b appears to have a low background level of expression at 22°C.

In light of this difference in regulation, a comparison of the promoters of *COR*15a and *COR*15b raises some questions (Figure 3.10). Both contain at least one g-box, an element that has been found in plant genes responding to a variety of environmental stresses (Daugherty et al., 1994), and at least one DRE-like element. The DRE has been shown to be involved in both cold- and drought-regulated induction (Yamaguchi-Shinozaki and Shinozaki, 1994). Finally, a 14-base sequence, designated the "H' repeat" is found twice in the promoter of *COR*15b and once in the promoter of *COR*15a. It is not known if this repeat affects inducibility. It has already been demonstrated that the drought-inducibility of *COR*15a is promoter-based (Baker et al., 1994). Is the same true of *COR*15b? If so, why does the DRE-like element not cause drought inducibility? This second question, while important, is beyond the scope of this study.

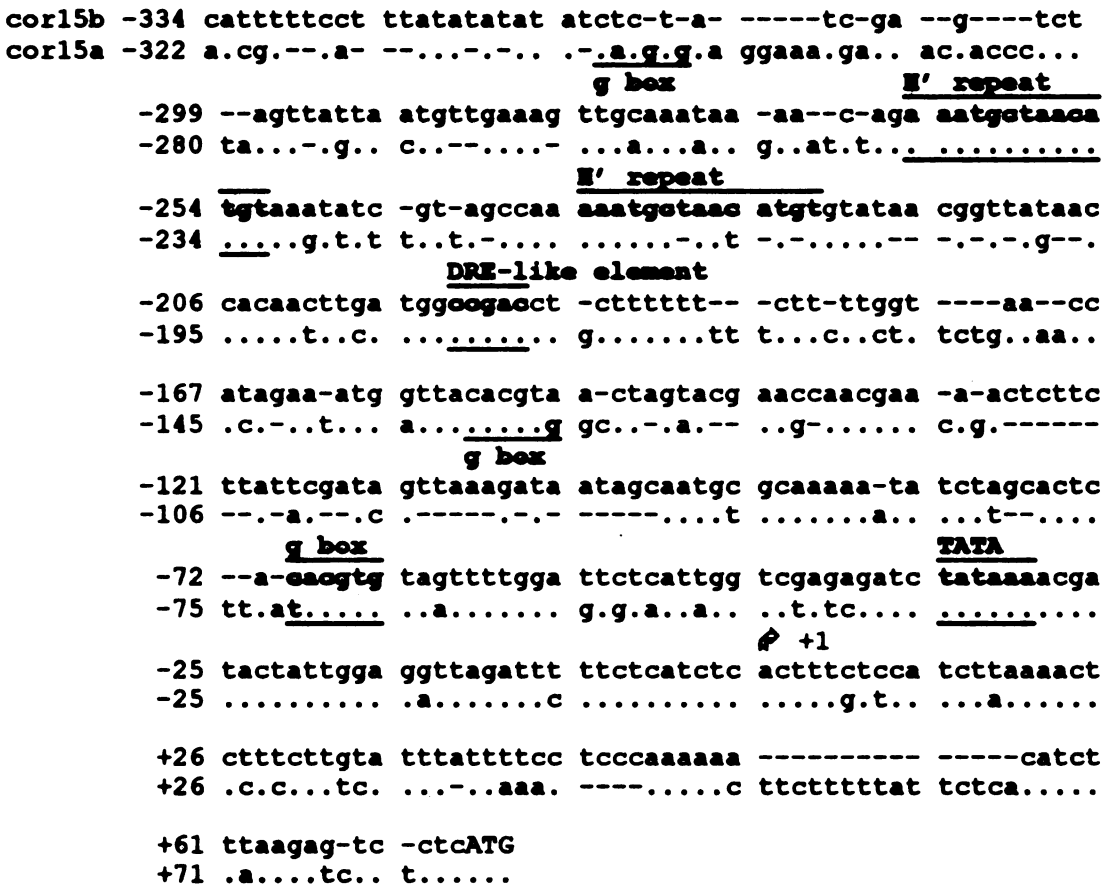


Figure 3.10 Comparison of the promoters of *COR*15a and *COR*15b. Dots represent identical bases and dashes mark gaps introduced for optimal alignment. Promoter elements are bold and indicated by an overline or underline for *COR*15a and *COR*15b, respectively. The start of transcription is indicated by an arrow, and the capitalized ATG is the translation start codon.

As an initial assessment of whether the *COR*15b promoter is low temperature or drought-responsive, two constructs were prepared and introduced into *Arabidopsis*. A 1,744 base pair fragment, including the *COR*15b promoter and part of the *COR*15b coding region (-1212 to +535 relative to the start of transcription), and a 300 base pair fragment of the *COR*15b promoter (-226 to +75) were each fused to the coding region of GUS to create 101.BS and 101Nla3, respectively. Construct 101.1BS was intended to be a translational fusion, but sequencing revealed that there had been a two base pair deletion during the cloning process, so the *COR*15b and GUS messages are out of frame.

Northern analysis of three lines of plants transgenic for 101.1BS revealed that the introduced gene is strongly responsive to low temperature, but not to drought (Figure 3.11). This demonstrates that induction by low temperature and the lack of drought inducibility are controlled by the promoter, the first exon, the intron or some combination of these. Two of the three lines of 101Nla3 tested showed the same responsiveness to low temperature and lack of drought-inducibility, but the third showed slight drought induction (Figure 3.11). The reason for this is unclear.

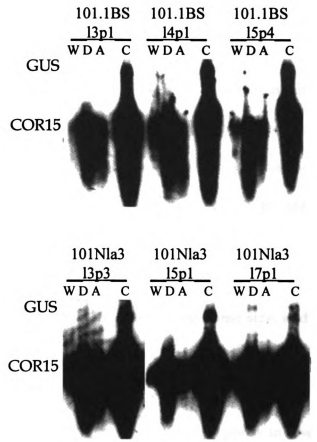


Figure 3.11 Regulation of *COR*15b promoter constructs. 101.1BS consists of a large fragment of the *COR*15b promoter (-1212 to +535 relative to the start of transcription) fused to the GUS coding region, and 101Nla3 is the smaller *COR*15b promoter fragment (-226 to +75 relative to transcription start) fused to GUS. Blots of three lines of each are shown here, probed for both GUS and endogenous *COR*15. Plants were grown at warm (normal) conditions (22°C, W), drought-stressed (D), treated with ABA (A), or cold-treated (2.5°C, C).

Discussion

Analysis of genomic Southerns of *Arabidopsis* has revealed that *COR*15 is a gene family. Southern and sequence analysis of a genomic clone containing *COR*15a has further shown that the second member of this family is situated less than 1 kb downstream of the first. If there are other members of this family, their sequences are too highly divergent for them to be detected with the pHH67 (Hajela, 1990) probe at low stringency. Sequence analysis and comparison then indicated that these two members of the family share many characteristics. They share high sequence identity, are in the same orientation, are predicted to have an intron in the same place, and the proteins encoded by both appear to have a chloroplast signal peptide, to be about the same size, and to have the same pI and structure.

The roles of all of the *COR* proteins in the response to low temperature are currently unclear. It has been shown that warm-grown plants expressing transgenic *COR*15a are less susceptible to freezing damage, as measured by chlorophyll fluorescence and by protoplast survival, than are wildtype plants, but the mechanism for this protection is unknown. COR15am and COR15bm show high amino acid sequence identity with the predicted mature forms of BN115, BN19 and BN26 from *Brassica napus* (Weretilnyk et al., 1993), with the predicted protein product of a cold-responsive Brassica oleracea mRNA

(accession number U16751; Jo H, Lee C and Sohn U, unpublished data) and low identity with repeat regions characteristic of Group 3 LEA proteins (Dure, 1993). Within group 3 LEA proteins, there are two subclasses. The degenerate repeats of COR15b show greater similarity with the repeats of group 3 LEA (I) (Dure, 1993), but COR15b is acidic, as are proteins classified as group 3 LEA (II) (Curry and Walker-Simmons, 1993). Unfortunately, neither structure nor function is known for any of these groups of proteins.

In analyzing the function of a protein, it would be helpful to have a model of its structure, so as to know which residues could be positioned to interact with a substrate, a solvent or another member of a complex. Specific residues could then be mutagenized to determine which are required for activity, and comparisons to proteins of similar structure could be made. It is also helpful to know the structure so that mutants can be designed to not disrupt protein folding. Mutations are typically more permissible at the solvent-exposed surface and more likely to relate to protein function than residues buried in the protein core. Also, mutations, including insertions and deletions, are typically only allowed in turn- or loop- forming regions, rather than within secondary structures.

Three secondary structure prediction programs have all indicated that COR15am and COR15bm are highly helical. In the process of making these predictions, it

was discovered that part of a protein of known structure, a subdomain of beta-lactamase, aligned with COR15am with high enough amino acid sequence identity to suggest the use of a knowledge-based modelling approach in developing a model for COR15bm tertiary structure.

The model was constructed with particular caution because the sequence identity between COR15am or COR15bm and beta-lactamase is at the lower limit for knowledge-based modelling (Sander and Schnieder, 1991) and because it would be unexpected for COR15am or COR15bm to function in antibiotic cleavage, as does beta-lactamase. However, the model does agree with the secondary structure prediction of PHDsec and does conform to the features of known protein structures. Namely, hydrophobic residues are buried and hydrophilic residues are exposed to the solvent. The two insertions occurring in COR15bm relative to beta-lactamase are in loops between helices, and thus would not disrupt the overall structure of this domain. Finally, the model shows a globular fold of amphipathic alpha helices which comprise one compact domain of the beta-lactamase structure. This domain has very few interactions with the other domain so it is expected that the domain corresponding to COR15bm could fold similarly in solution in the absence of the second domain. However, the sidechains in one of the surface loops are probably more tightly packed in COR15bm than in the model, since the predicted structure has a

tunnel through the periphery, which is unexpected for a globular protein and is not found in the beta-lactamase from which the model was derived.

More support for the model comes from the observation that the COR15m proteins are predicted be highly helical and amphipathic and are known to be very soluble in aqueous solution. It is reasonable to think that these proteins either multimerize or fold to shield the hydrophobic face of the helix or helices from the surrounding solvent. It is not yet known whether they multimerize. If they do not multimerize, then it is quite likely that they fold, possibly as predicted by this model which does generally have the hydrophobic residues buried.

It is unlikely that COR15m is a beta-lactamase. It contains three of the residues involved in catalysis (Ser, Lys and Gly) in the same general region of the structure as beta-lactamase, but corresponds to less than half of the beta-lactamase molecule and lacks the domain contributing the other half of the binding site. It does, however, show some structural similarity to another class of molecules, the lipoproteins, which bind lipids for transport through the cytoplasm. When analyzed by threading, the mature forms of both COR15a and COR15b were favorably compared to apolipophorin III from the African locust *Locusta migratoria* (Breiter et al., 1991). While this comparison was not useful for

modelling, it provoked some thought regarding the possible function of COR15am and COR15bm.

The apolipophorins in insects (Breiter et al., 1991) and plant nonspecific lipid transfer proteins (nsLTPs) (Heinemann et al., 1996) are composed of small bundles of amphipathic helices that are thought to spread their hydrophobic core over lipids, thus solubilizing them for transport through the aqueous cytosol (Breiter et al., 1991). Mature COR15m is similar in size to plant nsLTPs, but has a low pI rather than the high pI characteristic of these proteins, and lacks their eight conserved cysteines. Thus, the COR15m proteins are not plant nsLTPs, but may belong to a class of lipid-binding proteins that has not yet been described. Indeed, acidic phospholipid exchange proteins have been reported in castor bean (Tanaka and Yamada, 1982) and spinach (Kader et al., 1984), although their sequences are not known. Also, it has been found that the common structural feature of multiple amphipathic helices is more important for the lipid transfer function of insect apolipoproteins than amino acid identity (Smith et al., 1994). This lends credence to the idea that COR15m could have lipid transfer activity, even though its sequence differs substantially from that of other lipid transfer proteins.

Because membranes are a principal site of damage during a freeze-thaw cycle and because membrane composition changes during cold acclimation, a lipid

transfer function for COR15m would not be unexpected. The presence of COR15am, and possibly COR15bm, in the chloroplast would also be consistent with a potential role in lipid transport, since the synthesis of glycerolipids originates in the chloroplast (Roughan et al., 1980). *In vitro* assays such as those described by Miquel et al. (1987) could be utilized to determine whether the COR15m proteins are able to facilitate the transport of lipids. If such activity could be detected, then the model could be used further to direct a mutational analysis of amino acid residues in order to clarify which residues are necessary for this function.

The similarity of the two genes and the proteins they are thought to encode suggests that either more of the gene product is needed than can be produced from one copy of the gene or that the two function in different times, places or ways. None of these possibilities have been ruled out, although it has been seen that the principal difference between the two genes is the way they are regulated. *COR*15b, as studied by promoter/GUS fusions and by a probe specific for its transcript, has a low level of expression in warm-grown plants, while *COR*15a does not (Baker et al., 1994). More notably, however, *COR*15a is clearly inducible by drought stress, while *COR*15b is induced, if at all, to a substantially lower level.

A construct containing a little over a kilobase of the *COR*15b promoter, the first exon and the intron of *COR*15b, 101.1BS, was inducible in transgenic plants by low temperature, but not by drought, indicating that low temperature induction and the lack of drought induction are regulated by the promoter or by the first half of the transcript. A second construct containing about 300 bases of promoter sequence and no coding sequence, 101Nla3, gave more variable results. All of the lines tested were inducible by low temperature, indicating that this is controlled at the promoter level, but one of them also showed induction in drought-stressed plants.

With only three reliable samples to examine, it is difficult to know which is aberrant. It is possible that an element which suppresses drought-inducible expression was not included in this small construct and that all three lines should have shown drought induction. It is also possible that all three lines should not have shown drought-induction, similar to the results from 101.1BS. Alternatively, two of the lines could be showing the effects of trans-inactivation. This seems unlikely, since the constructs are still inducible by low temperature. Whatever the case, the results seem to reflect the effect of the position of the transgenes in the genome. Thus, more work needs to be done to determine why the endogenous *COR*15b gene is not responsive to drought stress, but the current evidence suggests that the lack of inducibility is promoter-based. More transgenic lines have been generated, which could be tested, although the

possibility that the DRE-like element present in the *COR*15b promoter is primarily responsible for induction due to low temperature but not drought is enticing. Answering the question of whether or not this element responds to both stimuli could ultimately be more informative than trying to unravel these transgenics.

Why does the DRE-like element in the promoter of *COR*15b not cause *COR*15b RNA levels to be inducible by drought stress? It is possible that message is induced but also degraded more quickly under drought conditions than low temperature. A construct containing the coding region of *COR*15b driven by the CaMV 35S promoter has been made to test this possibility, but the answer is not yet known.

If, however, the promoter of *COR*15b is not responsive to drought, there are two reasonable possibilities regarding the role of the DRE-like element. Bases surrounding the CCGAC core could affect its responsiveness or it could require interaction with a second promoter element. It has already been shown that the core g-box (ACGT) is not specific to ABA inducibility, but is required for induction under numerous stresses (Foster et al., 1994). G-boxes with slightly differing surrounding sequences (Foster et al., 1994) and in combination with differing promoter elements (Kawagoe et al., 1994, Shen and Ho, 1995 and Hatton et al., 1995) are instrumental in responses to different stresses. Thus it is

possible that the nucleotides surrounding the DRE-like element in the *COR*15b promoter cause this promoter element to be non-responsive to drought stress. This would be congruent with the observation that a *COR*15a promoter deletion construct which includes only the DRE-like element identical to the *COR*15b DRE-like element in the context of the native promoter is not very drought-responsive, while *COR*15a promoter deletion constructs including two more DRE-like elements, which have slightly differing surrounding sequences, are more strongly responsive to drought (Baker et al., 1994). In other experiments, two DRE-like elements found in the *rd29A* (*COR78*) promoter, which contain the DRE CCGAC core sequence, did not show drought-inducibility in tobacco (Yamaguchi-Shinozaki and Shinozaki, 1994).

Alternatively, cooperation with a second *cis*-acting element not present in this promoter may be required. The results of Yamaguchi-Shinozaki and Shinozaki (1994) suggest that a single DRE derived from *rd29A* (*COR78*) is sufficient for drought-inducible expression, since an 83 bp fragment containing the upstream DRE or a 53 bp fragment containing the downstream DRE fused to a minimal promoter derived from the same gene and to GUS gives drought-inducible GUS activity. Further, five copies of a 25 bp fragment containing the proximal DRE, when fused to the minimal promoter of *rd29A* (*COR78*) driving GUS, conferred drought-inducibility to GUS in transgenic tobacco (Yamaguchi-Shinozaki and Shinozaki, 1994). It is possible that there is a second promoter element contained

within each of these fragments, although it is not likely that the 25-base fragment would contain another element. However, the presence of multiple copies of the same element might make up for the lack of a coupling element. It is also possible that the minimal promoter used contains the required second element or that the function of these promoter elements is different in tobacco than in *Arabidopsis*. *Arabidopsis* is chilling tolerant and is able to acclimate to freezing temperatures, while tobacco is chilling sensitive and unable to acclimate to freezing.

In order to test whether the DRE-like element from *COR*15a or *COR*15b is inducible by both cold and drought stress or only by cold stress, experiments comparing the inducibility of constructs containing an oligonucleotide of the DRE from *COR*78, or the DRE-like elements of *COR*15a or *COR*15b fused to a minimal promoter and a reporter should be carried out in *Arabidopsis*. Constructs containing other promoter elements along with the DRE-like elements or containing multiple copies of the DRE-like element could give insight into whether the elements are functioning synergistically.

Both *COR*15a and *COR*15b appear to be inducible by ABA in gene-specific northern analysis. This result was called into question when the *COR*15b promoter constructs were not seen to be induced by ABA (Figure 3.11), but the ABA induction of the endogenous gene was weak enough that it is not clear that

induction would have been detectable. The gene specific northern blots need to be repeated, since only one ABA-induced RNA sample was examined initially, and the transgenics must be tested with better ABA induction before it can be unequivocally determined whether *COR*15b is responsive to ABA.

The phenomenon of *COR* genes existing as small gene families, members of which are within three kilobases of each other on a chromosome and are differentially regulated, has now been seen for all four *COR* gene families in *Arabidopsis* (Figure 3.12). *COR*6.6 and *kin1* are related to each other essentially the same way *COR*15a and *COR*15b are related to each other (Kurkela and Borg-Franck, 1992). Both show induction by low temperature and ABA, but only *COR*6.6 is strongly drought-responsive and only *COR*6.6 is expressed at low levels in control plants (Kurkela and Borg-Franck, 1992). However, Wang et al. (1995) used promoter fusions of these two genes to show that the *kin1* and *COR*6.6 promoters both responded to drought stress. This reason for this difference has not yet been reported. *COR*78 and *rd29B* differ in that the former is more strongly cold-responsive, while the latter is more strongly drought-regulated (Yamaguchi-Shinozaki and Shinozaki, 1993), and *COR*47 is more strongly responsive to drought and ABA than is *lti45* (also called *lti29*) (Welin et al., 1994 and Welin et al., 1995).

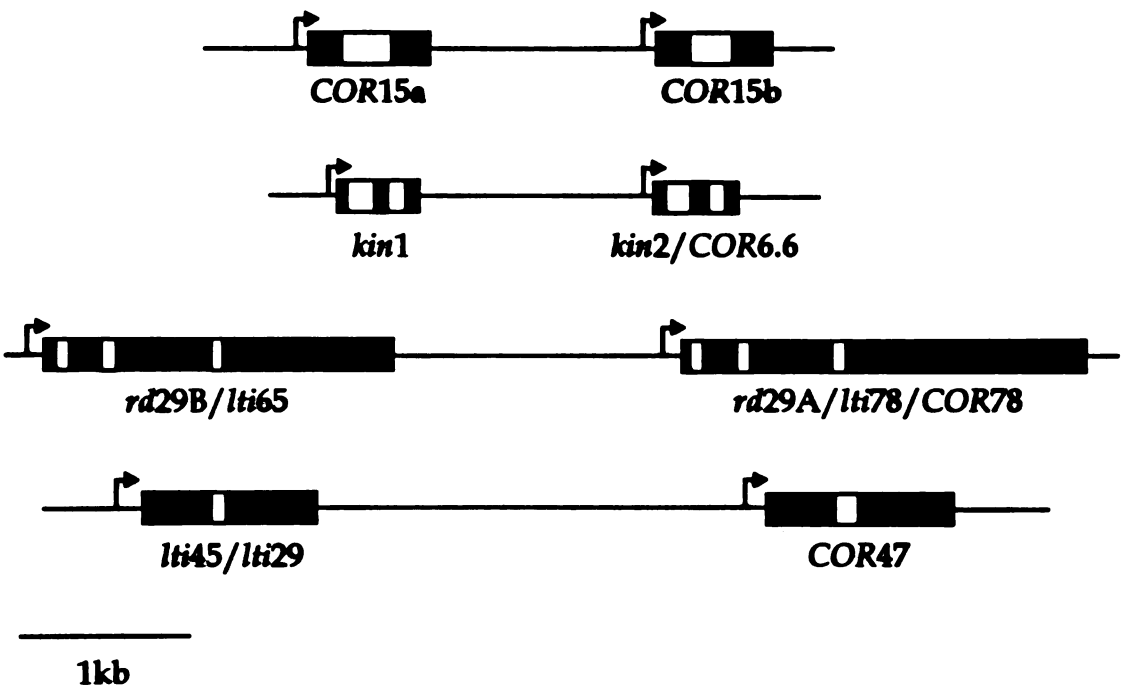


Figure 3.12 Genomic organization of *COR* gene families. Coding regions are shown as filled boxes, while introns are depicted as open boxes. Transcription start sites are indicated by arrows.

The members of each of the four gene families share 49% (*COR78* and *rd29B*), 67% (*COR47* and *lti45*) and 91% (*kin1* and *COR6.6*) amino acid identity (Yamaguchi-Shinozaki and Shinozaki, 1993, Welin et al., 1995, Kurkela and Borg-Franck, 1992 and Wang et al., 1995). These high amounts of identity suggest that the members of a given gene family may have identical or related functions. Perhaps the small differences among members of a gene family make one member better suited for protection of the plant under drought stress conditions without the addition of low temperature.

The shared spatial organization of the members of these families is also intriguing. It is possible that their proximity to each other is advantageous in that the opening of the chromatin in one region allows the transcription of two cold-inducible genes for the price of one. With this in mind, the question of where the *COR* gene families are in relation to each other on the chromosome becomes more interesting. Perhaps there is one loop of chromatin that contains all four families and becomes available for active transcription when the cell is exposed to low temperature. Alternatively, perhaps *COR* gene families are scattered in the genome, but are near other, less highly-expressed cold-inducible genes.

Materials and Methods

Sequencing and sequence analysis

Sequencing of the region of the genomic subclone corresponding to the *COR*15a cDNA pHH67 (Hajela et al., 1990), pBlue67, and of the junction between the GUS gene and *COR*15b promoter of 101Nla3 and 101.1BS was performed using the method of Sanger et al. (1977), using double-stranded DNA templates as described previously (Lin and Thomashow, 1992). Deletions of the genomic DNA insert were generated by digestion with exonuclease III and mung bean nuclease (Henikoff, 1987). Automated fluorescent sequencing of *COR*15b was performed by the MSU-DOE-PRL Plant Biochemistry Facility using the ABI Catalyst 800 for Taq cycle sequencing and the 373A Sequencer for the analysis of products.

Nucleotide sequence analysis was carried out using MacVector™ 3.5 (International Biotechnologies, Inc.), and protein analysis was done using Protean (DNASTAR), PHDsec and PHDthreader (www address: http://www.embl-heidelberg.de/predictprotein/phd_pred.html). Protein modelling was performed using InsightII (version 95.0).

Plant genomic DNA extraction and Southern analysis

DNA extraction for the determination of the number of members of the *COR*15 gene family was performed as described by Rogers and Bendich (1988) except that the mixture of plant tissue and extraction buffer was incubated at 68°C for 30 minutes before chloroform extraction and the CTAB precipitation was left until flocculant precipitate was seen. This sometimes took three or four days, and the same procedure used with the same CTAB a year or so later gave varying yields, so DNA preps used in analyzing the *COR*15b promoter/GUS transgene inserts were performed according to Tai and Tanksley (1991) as modified by Doug Dahlbeck. Up to 0.15 g leaf tissue was frozen in a microcentrifuge tube in liquid nitrogen, crushed with a miniature pestle and incubated at 65°C with 0.7 ml preheated extraction buffer (100 mM Tris-HCl pH 8, 50 mM EDTA pH 8, 500 mM NaCl, 1.25% SDS, 8.3 mN NaOH, 0.38% sodium bisulfite) for 10 minutes. Then 0.22 ml 5 M potassium acetate was added, the samples were incubated on ice for 40 minutes and microcentrifuged at 13,000 rpm for 3 minutes. The supernatant was filtered through a piece of Kimwipe® in a 1 ml pipetman tip into another microcentrifuge tube and precipitated with 0.7 volume of isopropanol. The pellet was resuspended in 300 µl T_5E (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8), allowed to resuspend at 65°C for 5 minutes, mixed with 150 µl 7.4 M ammonium acetate and microcentrifuged for 3 minutes. The supernatant was then transferred to a new tube and DNA was precipitated with 330 µl isopropanol. The pellet was washed twice with 70% EtOH, allowed

to resuspend at 65°C for 5 minutes in 100 µl T_5E and precipitated with 10 µl sodium acetate and 75 µl isopropanol. Finally, the pellet was washed twice with 70% EtOH, dried and resuspended in 25 µl sterile double distilled water.

Total *Arabidopsis* DNA or plasmid DNA was digested with various restriction enzymes and fractionated by agarose gel electrophoresis using TAE, which is 40 mM Tris-acetate/1 mM EDTA (Sambrook et al., 1989). Southern blots were then prepared using standard methods (Sambrook et al., 1989) and the DNA was UV cross-linked (Stratalinker, Stratagene) to the filters.

Standard genomic Southern blots were prehybridized and hybridized in a Robbins Scientific Model 400 hybridization oven using the nonfat dry milk method of Johnson et al. (1984) and washed using standard conditions (Sambrook et al., 1989), except that all washes were done with 0.1% SDS/0.1x SSC. High stringency blots were prehybridized in 2x SSC/0.25% nonfat dry milk/0.5% SDS at 55°C for three hours, hybridized overnight in the same solution at 68°C, and washed in 0.1x SSC/0.1% SDS. The first four 20-minute washes were at room temperature and the last two were at 68°C. Low stringency blots were prehybridized in 6x SSC/0.25% nonfat dry milk/0.5% SDS for three hours at 55°C, hybridized in the same solution overnight at 64°C and washed four times, for 20 minutes each time, at room temperature and twice, for 20 minutes each time at 50-60°C in 6x SSC/0.1% SDS. Gel-isolated fragments

labeled with ^{32}P by random priming (Feinberg and Vogelstein, 1983) were used as probes. A *Sac*I/*Hind*III fragment of pBI101.3 was used to detect GUS inserts, and an *Eco* RI fragment of pHH67 was used to visualize *COR*15a and *COR*15b sequences.

Plant growth and stress treatment

Arabidopsis thaliana (L.) Heyn. ecotype RLD or Columbia was grown in controlled environment chambers at 22°C with a 24 hour photoperiod (about 120 μmol m^{-2} s^{-1}) as previously described (Gilmour et al., 1988). Humidity was not controlled. After about three weeks of growth, plants were either harvested or subjected to one of several stress treatments. Plants were cold-acclimated by placing pots in a 2.5°C cold room under constant cool-white fluorescent light (45-55 μmol $m^{-2}s^{-1}$) for four days. Drought stress was administered by excising the shoots from the pots, placing them on 15 cm Whatman No. 1 filter paper in 150 mm x 15 mm plastic petri plates with the lids open for 45 minutes. The lids were then placed on the plates and plants were incubated for four to six hours under fluorescent lights (about 50 μmol $m^{-2}s^{-1}$). Water loss was quantified by dividing the weight of the drought-stressed plants by their initial fresh weight and multiplying by 100. Plants lost from 18% to 23% of their initial fresh weight. The ABA treatment of plants used in the gene-specific northern experiments consisted of spraying plants to runoff with 50 μM ABA (mixed isomer, Sigma) in 0.02% polyoxyethylene-sorbitan monolaurate 20 (Tween 20), covering the pots

with Saran wrap to slow evaporation and incubating the plants for 4h at 22°C. All other ABA treatments were performed by placing excised plants into a 150 mm x 15 mm petri plate with 75 ml of 100 μM ABA (mixed isomers, Sigma) and one drop (10-20 μl) Tween-20, covering them with a piece of 15 cm Whatman No. 1 filter paper to keep them submerged, placing the lid on the plate and incubating them for 4 to 6 hours under fluorescent lights (about 50 μmol $m^{-2}s^{-1}$). Water-soaked plants underwent this same treatment, except that the ABA was omitted.

RNA extraction and fractionation

Total RNA was extracted essentially as described (Gilmour et al., 1988) with a few modifications. Frozen pulverized tissue was extracted with equal volumes of phenol/chloroform/isoamyl alcohol (25/24/1, v/v/v) and extraction buffer (1% w/v triisopropylnaphthelene sulfonic acid, 6% w/v p-aminosalicylic acid, 100 μM Tris-HCl pH 7.6, 50 mM EGTA, 125 mM NaCl, 1% w/v SDS, 10 mM DTT) on ice. This was further homogenized in the tube using a Tekmar® Tissumizer, centrifuged (10,000 rpm in the SA600 rotor), and the supernatant extracted again with phenol/chloroform/isoamyl alcohol. This second set of tubes and all subsequent solutions except ethanol were made free of RNase by treatment with DEPC (Sambrook et al., 1989). Nucleic acids were precipitated with cold 95% ethanol, resuspended in 1 ml double distilled water, transferred to a microcentrifuge tube and precipitated on ice for an hour with 1/4 volume 10

M LiCl. The pellets were precipitated again with 95% EtOH and resuspended in 200 µl water. The OD_{260} was measured to estimate the concentration of RNA.

For fractionation, RNA (5 to 40 µg) was dried down and resuspended in formaldehyde loading buffer containing EtBr (about 100 ng/ml), incubated at 68°C for 10-15 minutes and fractionated on denaturing formaldehyde agarose gels (Sambrook et al., 1989). The RNA was then transferred (Sambrook et al., 1989) to Magna NT membranes (MSI) using 10x SSPE made according to instructions provided by Schleicher and Schuell. This recipe for SSPE is slightly different from that described by Sambrook et al. (1989). RNA was UV cross-linked (Stratalinker, Stratagene) to the filters.

Primer extension

Primer extension, performed as described by Horvath et al. (1993) and using the oligonucleotide 5'-CAGCTCCTGAGAAAGACATCGC-3', was used to determine the transcription initiation site of *COR* in cold-treated plants.

Oligonucleotide northern analysis

Oligonucleotide northerns were done using the gene specific oligonucleotides 5'TGCTTGAAATTAACTGATTAGGTAAGACCC3' (*COR*15a) and 5'GAACATGACTACATGAGTGGTTGAATCAGG3' (*COR*15b), which were end-labeled as described by Zeff et al., (1987) with minor modifications (Baker,

1990). First, 20 µl of 200 µCi γ-^{32}P ATP (>5000 Ci/mmol) were dried and resuspended in 1 µl oligonucleotide (100 ng/µl), 1 µl T4 polynucleotide kinase, 1 µl 1 M Tris-HCl pH 8, 1 µl 0.1M MgCl, 2.5 µl 0.02M DTT, and 3.5 µl double distilled autoclaved water. This reaction mix was incubated for 30 minutes at 37°C, then inactivated at 65°C for five minutes and brought to 100 µl with TE (10 mM Tris-HCl pH 8, 1 mM EDTA). Unincorporated nucleotides were removed using a Sephadex G-25 spun column (Sambrook, 1989).

Hybridization was performed as described (Baker, 1990), except that the temperature used was 63°C instead of 42°C. This temperature had been empirically shown (Sambrook et al., 1989) to give gene-specific binding of the two probes. Filters were incubated in prehybridization solution (5x SSC, 10x Denhardt's solution, 7% SDS, 100 µg sheared, denatured fish sperm DNA) for four hours at 63°C, then in hybridization solution (5x SSC, 1% SDS, 10% dextran sulfate, 100 µg sheared, denatured fish sperm DNA) at 63°C overnight. The filters were washed once in 5x SSC/5% SDS/5x Denhardt's solution at 63°C for 30 minutes, once in 3x SSC/2% SDS at 63°C for 45 minutes and once in 2x SSC for 10 minutes at room temperature.

Cloning of *COR*15b promoter constructs

In the construction of 101Nla3, an *Nla*III fragment containing approximately 300 bp of the *COR*15b promoter (-226 to + 75 relative to the transcription start site)

was ligated into the *Sph*I site of pUC19 (Yanisch-Perron et al., 1985), such that the 5′ end of the fragment, relative to the *COR*15b coding region, is nearest the *Hin*dIII site of pUC19. The insert was then cut out with *Hin*dIII and *Xba*I and cloned into pBI 101.1 (Jefferson, 1987) cut with *Hin*dIII and *Xba*I.

Plasmid 101.1BS was made by using the large fragment of DNA polymerase I (Klenow) to fill in the 5′ overhang of a *Sca*I/ *Hin*dIII fragment which contains approximately 1.5 kb of the promoter (-1212 to +535) and a about a third of the coding region of *COR*15b. This fragment was cloned into the *Sma*I site of pBluescript (SK-) such that the 5′ end relative to the coding region of *COR*15b is nearest the *Pst*I site of the pBluescript polylinker. This new construct was then cut *Sal*I/*Bam*HI and the insert was gel-isolated and ligated into pBI101.1 (Jefferson, 1987), which had also been cut with *Sal*I and *Bam*HI. This should have put the chloroplast signal peptide of *COR*15b in frame with GUS.

Plant transformation

The pBI 101.1 recombinants were mobilized into *Agrobacterium tumefaciens* strain C58C1 (pMP90) (Koncz and Schell, 1986) or EHA105 (Hood et al., 1993) by triparental mating, using strain pRK2013 as a helper. The root transformation technique of Valvekens et al., (1988) with the modifications described by Baker et al., (1994) was used to transform *Arabidopsis thaliana* (L.) Heyn., ecotype RLD.

Northern hybridizations

For standard northern hybridizations, blots were prehybridized, hybridized and washed in a Robbins Scientific Model 400 hybridization oven according to standard methods (Ausubel et al., 1987) with some modifications. Blots were incubated with prehybridization solution (50% v/v formamide, 5x SSPE, 50 mM potassium phosphate pH 8, 5x Denhardt's solution, 0.5% SDS, 100 μg/ml sheared, denatured fish sperm DNA) at 42°C for three hours to overnight and in hybridization solution (50% formamide, 5x SSPE, 50 mM potassium phosphate pH 6.5, 1x Denhardt's solution, 0.5% SDS, 5% dextran sulfate, 100 μg/ml sheared denatured fish sperm DNA) at 42°C overnight. Blots were rinsed and then given two 30-minute washes at room temperature in 2x SSPE/0.5% SDS, then washed two to three times, for 15 minutes each time, with 0.1x SSPE/0.5% SDS at 50°C.

Probes were made from gel-isolated fragments labeled with ^{32}P by random priming (Feinberg and Vogelstein, 1983). An *Eco*RI fragment from pHH67 was used to visualize *COR*15 message, and a *Sac*I/*Hin*d III fragment from pBI101.3 (Jefferson, 1987) served as a GUS probe.

Literature Cited

Abola EE, Bernstein FC, Bryant SH, Koetzle TF, Weng J (1987) Protein Data Bank. In, FH Allen, G Bergerhoff, R Sievers, eds, Crystallographic Databases-information content, software systems, scientific applications. Data Commission of the International Union of Crystallography, Bonn, Cambridge, Chester, pp107-132

Altschul AF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J. Mol. Biol. 215:403-410

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current Protocols in Plant Molecular Biology. Greene Publishing Associates and Wiley Interscience, New York, Chichester, Brisbane, Toronto, Singapore

Baker SS (1990) Cloning, identification and characterization of fertilization associated genes in *Pinus strobus* (L.) ovules. Ph. D. thesis, Ohio State University

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. 24: 701-713

Breiter DR, Kanost MR, Benning MM, Wesenberg G, Law JH, Wells MA, Rayment I, Holden HM (1991) Molecular structure of an apolipoprotein determined at 2.5-Å resolution. Biochemistry 30:603-608

Brown JWS (1986) A catalogue of splice junction and putative branch point sequences from plant introns. Nucl. Acids Res. 14:9549-9559

Chen CCH, Herzberg O (1992) Inhibition of beta-lactamase by clavulanate: trapped intermediates in cryocrystallographic studies. J. Mol. Biol. 224:1103-1113

Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47: 45-148

Curry J, Walker-Simmons MK (1993) Unusual sequence of group 3 (II) LEA mRNA inducible by dehydration stress in wheat. Plant Mol. Biol. 21:907-912

Daughtery CJ, Rooney MF, Paul A-L, de Vetten N, Vega-Palas MA, Lu G, Gurley WB, Ferl RJ (1994) Environmental stress and gene regulation. In, EM Meyerowitz, CR Somerville, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 769-806

Dunnill P (1968) The use of helical net-diagrams to represent protein structures. Biophys. J. 8:865-875

Dure III L (1993) A repeating 11-mer amino acid motif and plant desiccation. The Plant Journal 3:363-369

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6-13

Foster R, Izawa T, Chua N-H (1994) Plant bZIP proteins gather at ACGT elements. FASEB J. 8:192-200

Franz G, Hatzopoulos P, Jones TJ, Krauss M, Sung ZR (1989) Molecular and genetic analysis of an embryonic gene, DC8, from *Daucus carota* L. Mol. Gen. Genet. 218:143-151

Garnier J, Osguthhorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of gobular proteins. J. Mol. Biol. 120:97-120

Gavel Y, von Heijne G (1990) A conserved cleavage-site motif in chloroplast transit peptides. FEBS Lett. 261:455-458

Gilmour SJ, Hajela RK, Thomashow MF (1988) Cold acclimation in *Arabidopsis thaliana*. Plant Physiol. 87:745-750

Hajela RK, Horvath DP, Gilmour SJ, Thomashow MF (1990) Molecular cloning and expression of *cor* (*co*ld-*r*egulated) genes in *Arabidopsis thaliana*. Plant Pysiol. 93:1246-1252

Hatton D, Sablowski R, Yung M-H, Smith C, Schuch W, Bevan M (1995) Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. The Plant Journ. 7:859-876

Heinemann B, Andersen KV, Nielsen PR, Bech LM, Poulsen FM (1996) Structure in solution of a four-helix lipid binding protein. Protein Sci. 5:13-23

Heinz DW, Baase WA, Dahlquist FW, Matthews BW (1993) How amino-acid insertions are allowed in an α-helix of T4 lysozyme. Nature 361:561-564

Henikoff S (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. Meth. Enzymol. 155:156-165

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Research 2:208-218

Horvath DP, McLarney BK, Thomashow MF (1993) Regulation of *Arabidopsis thaliana* L. (Heyn) *cor78* in response to low temperature. Plant Physiol. 103:1047-1053

Hsing YC, Chen Z, Shih M, Hsieh J, Chow T (1996) Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. Plant Mol. Biol. In press

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol Rep. 5:387-405

Johnson DA, Gautsch JW, Sportsman JR, Elder JH (1984) Improved technique utilizing non fat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose. Gene Anal. Tech. 1:3-8

Joshi CP (1987) An inspection of the domain between putative TATA box and translational start site in 79 plant genes. Nucl. Acids Res. 15:6643-6653

Kader JC, Julienne M, Vergnolle C (1984) Purification and characterization of a spinach leaf protein capable of transferring phospholipids from liposomes to mitochondria or chloroplasts. Eur. J. Biochem. 139:411-416

Kawagoe Y, Campbell BR, Murai N (1994) Synergism between CACGTG (g-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein *β-phaseolin* gene. The Plant Journ. 5:885-890

Koncz C, Schell J (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. Mol. Gen. Genet. 204:383-396

Knox JR, Moews PC (1991) Beta-lactamase of *Bacillus licheniformis* 749/C. Refinement at 2 Å resolution and analysis of hydration. J. Mol. Biol. 220:435-455

Kurkela S, Borg-Franck M (1992) Structure and expression of *kin2*, one of two cold- and ABA-induced genes of *Arabidopsis thaliana*. Plant Mol. Biol. 19:689-692

Lin C, Thomashow MF (1992) DNA sequence analysis of a complementary DNA for cold-regulated *Arabidopsis gene cor15* and characterization of the COR15 polypeptide. Plant Physiol. 99:519-525

Madgwick PJ, Waley SG (1987) Beta-lactamase I from *Bacillus cereus*. Structure and site-directed mutagenesis. Biochem. J. 248:657-662

Marqusee S, Baldwin RL (1987) Helix stabilization by Glu⁻--Lys⁺ salt bridges in short peptides of *de novo* design. PNAS 84:8898-8902

Miquel M, Block MA, Joyard J, Dorne A-J, Dubacq J-P, Kader J-C, Douce R (1987) Protein-mediated transfer of phosphatidylcholine from liposomes to spinach chloroplast envelope membranes. Biochimica et Biophysica Acta 937:219-228

Puupponen-Pimia R, Saloheimo M, Vasara T, Ra R, Gaugecz J, Kurten U, Knowles J, Keranen S, Kauppinen V (1993) Characterization of a *birch (Betula pendula Roth.)* embryonic gene, BP8. Plant Mol. Biol. 23:423-428

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues. In SB Gelvin, RA Schilperoort, eds., Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, Boston, London, pp A6/1-A6/11

Rost B (1995a) Fitting 1D predictions into 3D structures. In, H Bohr, S Brynak, eds, Protein Folds. A Distance-based Approach. CRC Press, Boca Raton, pp 132-151

Rost B (1995b) TOPITS: Threading one-dimensional predictions into three-dimensional structures. In, C Rawlings, D Clark, R Altman, L Hunter, T

Lengauer, S Wodak, eds, The third international conference on intelligent systems for molecular biology (ISMB). AAAI Press, Menlo Park CA, pp314-321

Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. J. Mol. Biol. 232:584--599

Rost B, Sander C (1994) Combining evolutionary information and neural networks to predict secondary structure. Proteins 19:55-77

Rothstein SJ, Lahners KN, Lotstein RJ, Carozzi NB, Jayne SM, Rice DA (1987) Promoter cassettes, antibiotic-resistance genes, and vectors for plant transformation. Gene 53:153-161

Roughan PG, Holland R, Slack CR (1980) The role of chloroplasts and microsomal fractions in polar lipid synthesis from [1-^{14}C acetate by cell-free preparations from spinach (*Spinacia oleracea*) leaves. Biochem. J. 188: 17-24

Sander C, Schneider R (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins 9:56-68

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A laboratory manual, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. PNAS 74:5463-5467

Schiffer M, Edmundson AB (1967) Use of helical wheels to represent the structure of proteins and identify segments with helical potentials. Biophys. J. 7:121-135

Shen Q, Ho T-H D (1995) Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a g-box and a novel *cis*-acting element. The Plant Cell 7:295-307

Smith AF, Owen LM, Strobel LM, Chen H, Kanost MR, Hanneman E, Wells MA (1994) Exchangeable apolipoproteins of insects share a common structural motif. J. Lipid Res. 35:1976-1984

Soll J, Alefsen H (1993) The protein import apparatus of chloroplasts. Physiol. Plant. 87:433-440

Strynadka NCJ Adachi H, Jensen SE, Johns K, Sielecki A, Betzel C, Sutoh K, James MNG (1992) Molecular structure of the acyl-enzyme intermediate in β-lactam hydrolysis at 1.7Å resolution. Nature 359:700-705

Tai TH, Tanksley SD (1991) A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. Plant Mol. Biol. Rep. 8:297-303

Tanaka T, Yamada M (1982) Properties of phospholipid exchange proteins from germinated castor bean endosperms. In, JFGM Wintermans, PJC Kuiper, eds., Biochemistry and metabolism of plant lipids. Elsevier Biomedical Press B. V., North-Holland, Amsterdam, pp 99-106

Valvekens D, Van Montagu M, Van Lijsebettens M (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. PNAS 85:5536-5540

Wang H, Datla R, Georges F, Loewen M, Cutler A (1995) Promoters from *kin1* and *cor6.6*, two homologous *Arabidopsis thaliana* genes: transcriptional regulation and gene expression induced by low temperature, ABA, osmoticum and dehydration. Plant Mol. Biol. 28:605-617

Welin BV, Olson A, Nylander M, Palva ET (1994) Characterization and differential expression of *dhn/lea/rab*-like genes during cold acclimation and drought stress in *Arabidopsis thaliana*. Plant Mol. Biol. 26:131-144

Welin BV, Olson A, Palva ET (1995) Structure and organization of two closely related low-temperature-induced *dhn/lea/rab*-like genes in *Arabidopsis thaliana* L. Heynh. Plant Mol. Biol. 29:391-395

Weretilnyk E, Orr W, White TC, Iu B, Singh J (1993) Characterization of three related low-temperature regulated cDNAs from winter *Brassica napus*. Plant Physiol. 101:107-177

Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of the expression of a desiccation-responsive *rd29* gene of *Arabidopsis thaliana* and the analysis of its promoter in transgenic plants. Mol. Gen. Genet. 236:331-340

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. The Plant Cell 6:251-264

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119

Zeff RA, Geier SS, Natherson SG (1986) Molecular loss variants of the murine major histocompatibility complex: nonexpression of H-2K antigen associated with marked reduction in H-2K mRNA as determined by oligonucleotide hybridization analysis. Immunology 137:1366-1370

Chapter 4:

Examination of the low temperature induction profiles of *COR* gene families

Summary

The low temperature induction profiles of the *COR*6.6, *COR*15, *COR*47 and *COR*78 gene families and of a GUS gene fused to the *COR*15a promoter were examined in transgenic *Arabidopsis*. In each case, expression was first seen to increase at about 16°C and message levels gradually rose as the temperature was lowered, reaching a maximum at about 8°C. This pattern was seen whether the plants had been shifted down two degrees at a time or had been dropped from growth temperature to the temperature of interest, suggesting that induction was due to the low temperature *per se*, rather than to a given change in temperature. The induction profiles of the *COR* gene families were also examined in non-transgenic *fad*2-3 and *fad6* mutants to see if decreased levels of polyunsaturated fatty acids in the plasmalemma or chloroplast membranes, respectively, would affect low temperature regulation of the *COR* genes. The induction profiles were

indistinguishable from those of the non-mutant plants, thus failing to support the idea that the lipid composition of membranes influences the signal transduction pathway at which these *COR* genes are induced.

Introduction

The examination of the regulation of the *COR*78, *COR*47, *COR*15 and *COR*6.6 gene families with respect to a variety of stresses has shown that at least one member of each family is inducible under conditions of low temperature, drought and the application of ABA (Hajela et al., 1990). For several of these genes, it has been shown that this induction is at the promoter level (Baker et al., 1994, Wang et al., 1995, Horvath et al., 1993, Yamaguchi-Shinozaki and Shinozaki, 1993). However, at the start of this work, experiments looking more specifically at the temperature of induction or at the amount of induction at any given temperature were only preliminary (Thomashow et al., 1990). A more complete characterization of low temperature induction was undertaken to extend the understanding of *COR* gene induction and to provide a foundation for speculation regarding possible mechanisms of signal transduction.

First of all, it was not known if all of these *COR* genes have the same temperature induction profile. If there is only one low temperature signal

operative, then it would be expected that the *COR* genes would all respond similarly, but a difference in profile might suggest the existence of multiple signals. Also unknown was the temperature or change in temperature required for induction. Induction is clear at 4°C, but the behavior of the gene families between 22°C and 4°C had not been examined. Indeed, the final temperature might not be the determining factor. Los and colleagues (1993) have shown that the low temperature-inducible gene *desA* of *Synechocystis* PCC6803 is responsive to a change in temperature of greater than 5°C, rather than to a specific temperature.

Subsequent work from this group demonstrated that *desA* induction is a factor of membrane fluidity (Vigh et al., 1993), which raised the question of whether *Arabidopsis* mutants deficient in the saturation levels of membrane fatty acids would show the same temperature induction profile as wildtype *Arabidopsis*. Plants containing mutations *fad* 2-3, in which levels of polyunsaturated fatty acids are increased in extra-chloroplastic membrane lipids (Miquel and Browse, 1992), and *fad 6*, which have an increased levels of polyunsaturated fatty acids in chloroplast membrane lipids (Browse et al., 1989), were employed to examine this question.

Finally, although it was known that induction at 4°C is regulated at the promoter level, it was not clear whether this was true of regulation at other temperatures.

Results

In order to answer these questions, the temperature induction profiles of endogenous *COR* genes and of GUS message driven by the promoter of *COR*15a were determined. Plants containing the -900/+78 promoter fragment of *COR*15a fused to the GUS coding region (Baker et al., 1994) were grown under standard growth chamber conditions. Plants containing the *fad*2-3 or *fad*6 mutation were included in the same experiment. After three weeks of growth, the temperature was shifted two degrees a day from 16°C to 6°C. The plants received 24 hours of exposure to each temperature and a pot of each genotype was harvested immediately before the temperature was lowered again. Thus, the plants harvested at 6°C had been exposed to each of the temperatures in the series, culminating with 24 hours at 6°C (Figure 4.1).

Northern analysis of the RNA extracted from these samples, using probes specific for *COR* gene families (but not specific for individual members of those families) or GUS, did not show any clear difference among the temperature induction profiles of the *COR* gene families (data not shown) the transgene (Figure 4.2) or the *fad* mutants. Message levels were detectable over background at 16°C, gradually increased as the temperature was lowered, then appeared to level off at and below 8°C.

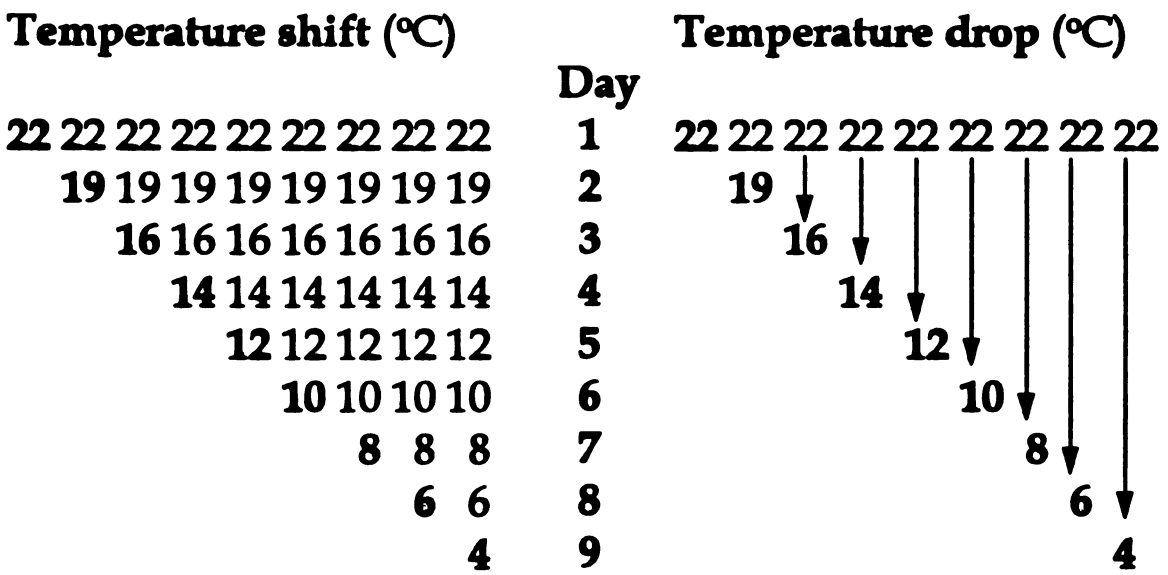


Figure 4.1 Two methods of lowering the temperature. For both methods, plants were grown for about three weeks at 22°C. In the temperature shift method, the temperature was lowered by two or three degrees at a time. A pot of plants was harvested after exposure to 24 hours of a given temperature, but had also experienced 24 hours at each of the higher temperatures in the series. In the temperature drop method, the plants were kept at 22°C and then taken directly to one of the lower temperatures in the series. Thus most of these plants experienced only one of the low temperatures in the series and experienced a drop in temperature of up to 18°C.

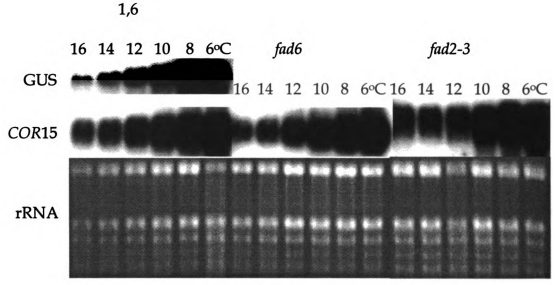


Figure 4.2 First temperature shift experiment. RNA isolated from 1,6 plants at the temperatures listed above the lanes is shown probed for GUS (GUS is fused to the *COR*15a promoter in these plants) and *COR*15. RNA isolated from fatty acid deficient mutants *fad6* and *fad2-3* at the given temperatures is only shown hybridized with a *COR*15 probe. Ethidium bromide stained rRNA is included to show differences in loading.

Upon repetition of the experiment, the plants were similarly grown, but the range of the temperatures was broadened to include 22°C (standard growth temperature), 19°C and 4°C. This was done in order to determine whether there was a gradual increase in the amount of message at any temperature below 22°C or whether 16°C was the highest temperature at which induction occurred.

In a concurrent experiment, designed to determine whether induction due to this gradual lowering of temperature would differ from that following a single drop in temperature, two sets of the plants containing the *COR*15a promoter fused to the GUS coding region were grown together for three weeks. At this point, one set of pots was shifted down in temperature as just described, while the second set was maintained at 22°C. As the first set reached each low temperature, a pot from this second set was incubated with it at that temperature for 24 hours and then harvested. As an example, beyond the initial three weeks of growth, the plants of a given pot from the first set of plants would have experienced 22°C, 19°C, 16°C, 14°C and 12°C, each for 24 hours, while corresponding plants from the second set would have experienced four days of 22°C and 24 hours of 12°C.

Northern analysis of this second set of experiments confirmed the results from the first and showed that whether the temperature was lowered by two or three degrees at a time or was dropped more dramatically, the level of induction at each temperature was the comparable (Figure 4.3). Message levels were first

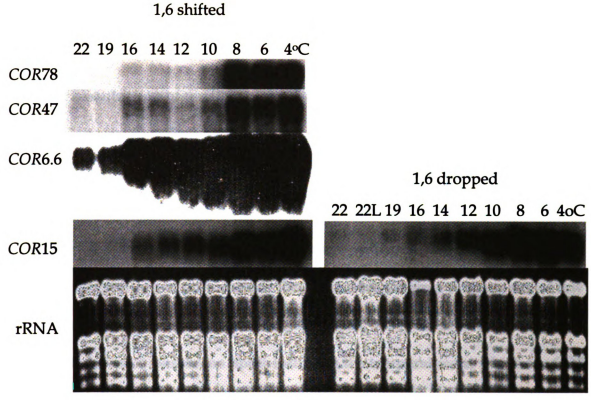


Figure 4.3 Second temperature shift experiment. Shown are northerns of 1,6 plants that were either shifted gradually down in temperature by two or three degrees at a time (1,6 shifted) or were dropped from 22°C to the temperature indicated above each lane (1,6 dropped). Lane 22L in the temperature drop series refers to a sample that was kept at 22°C for the duration of the experiment to control for the extended period of incubation at 22°C. The blot of 1,6 which was shifted down in temperature is shown here probed with *COR*78, *COR*47, *COR*15 and *COR*6.6. The rRNA is shown to indicate differences in loading. The *COR*6.6 probe used here gave very strong signal relative to the other probes, which is thought to be due to the probe itself, rather than to the message to which it hybridizes.

detectable over background at 16°C, gradually increased as the temperature was lowered, and again appeared to level off at and below 8°C.

Discussion

In the examination of possible mechanisms of signal transduction for the *COR* genes, several questions about the nature of the induction arise. What is the temperature induction profile of the *COR* genes? Do all the *COR* genes follow the same pattern with respect to low temperature? Are they induced by a specific temperature, or by the change in temperature? Is the temperature induction profile promoter-based? Since membranes are a primary site of low temperature injury and acclimation (Steponkus, 1984), do mutations that alter the saturation of membrane lipids influence the temperature at which the *COR* genes are induced?

Induction of message for all four of the gene families was first seen at 16°C, although it may increase at a slightly higher temperature, since 17°C and 18°C were not examined. The level of induction then gradually increased as the temperature was lowered, and appeared to reach a maximum at 8°C. This was true whether the plants reached a given temperature in increments or in a single drop, so the genes appear to be responding differently to different temperatures, irrespective of the change in temperature. Plants that received the three degree

shift from 22 to 19°C showed no induction, while those given a three degree shift from 19 to 16°C did show slight induction, strongly suggesting that the temperature and not the shift is responsible for induction.

It was also unclear whether the slight induction at the higher temperatures in the series (16, 14 and 12°C) was due to the temperature itself or to the rate at which message accumulates. Perhaps it takes longer to accumulate *COR* gene mRNA at 14°C than at 8°C, but they would reach the same level if given enough time. Comparison of the expression patterns of the temperature shift and temperature drop experiments suggested again that the temperature itself is responsible for the level of induction and that if induction is slower at higher temperatures, it still appears to reach saturation within 24 hours. Otherwise, the plants that had experienced each temperature in the series should have had a higher level of induction than those which had only received one low temperature.

The levels of transcript also suggest that there may be one major increase in message levels between 19 and 16°C and a second between 10 and 8°C. This may point toward two mechanisms of low temperature induction of COR genes: one which operates at 16-10°C and a second which involved in induction at and below 8°C. There could be two sensors of low temperature, two signal transduction pathways, or perhaps one transcriptional pathway and a second pathway which controls rates of mRNA degradation. Repetition of the

experiment and quantification of message levels at each temperature should help clarify this intriguing possibility. Further work could then be directed to examining signal transduction in the two temperature ranges separately.

Also, since the induction of GUS message driven by the *COR*15a promoter follows the same pattern as that of the endogenous *COR*15 gene family, it would appear that the temperature induction profile of *COR*15a is promoter-based. It has previously been shown that other aspects of low temperature induction are principally promoter-based in the *COR6.6* gene family (Wang et al., 1995) and for *COR78* (Horvath et al., 1993 and Yamaguchi-Shinozaki and Shinozaki, 1994), so it would be consistent for their promoters to behave like that of *COR*15a. Use of the promoters for the second member of the *COR78* gene family, *lti65* (Nordin et al., 1993), also called *rd29b* (Yamaguchi-Shinozaki and Shinozaki, 1994), and for members of the *COR47* gene family (Welin et al., 1995) in reporter gene fusions has not yet been described.

Finally, the lack of alteration in the induction pattern of the *fad* mutants suggests that bulk membrane saturation is not a significant factor in the induction of these *COR* genes. This fails to support a model in which a change in membrane fluidity is the signal prompting low temperature gene induction. However, it is still possible that the conformation of a specific membrane component or that a microenvironment formed in membranes exposed to low temperature might be

involved in low temperature signal transduction. It has been shown that calcium-selective ion channels in the plasmalemma of onion cells are sensitive to low temperatures in (Ding and Pickard, 1993).

Calcium has been strongly implicated in the regulation of cold-responsive genes in alfalfa (Monroy and Dhindsa, 1995). It has been shown that cold shock to 5°C or 0°C causes a rise in the intracellular concentration of calcium in tobacco (Knight et al., 1992). This could not explain the induction of the *COR* genes at 16°C unless calcium levels in *Arabidopsis* increase at a higher temperature than in tobacco, since cold shock to 10°C or higher in tobacco did not cause a rise in intracellular calcium (Knight et al., 1992). It is more encouraging to note that calcium influx was seen in alfalfa protoplasts at temperatures as high as 15°C (Monroy and Dhindsa, 1995). Cold shock with ice water elicited an increase in intracellular calcium levels in *Arabidopsis* (Knight et al., 1996), but calcium levels in *Arabidopsis* at other temperatures have not been reported. Work would need to be done to test whether temperature-responsive channels, particularly channels regulating calcium fluxes, could be found in *Arabidopsis* and to test what temperatures caused the effect.

The possibility remains that *COR* genes in *Arabidopsis* might be inducible upon return to normal growth temperatures following heat shock. It has been seen that the *COR* genes *BN115* and *BN28* in *Brassica napus* L. cv. Westar (a spring

cultivar), are induced when two-week-old plants are moved from a two-hour heat shock of 42°C to normal growth conditions for four hours (Krishna et al., 1995). However, when Weretilnyk et al. (1993) studied *BN115* in *Brassica napus* L. cv Jet neuf (a winter cultivar) they allowed three week old plants to recover at 20°C for two to nine hours after a 42°C heat shock before harvesting leaf tissue, and did not see induction. Perhaps the difference was due to the cultivar or to the age of the plants used. *Arabidopsis* will not survive 42°C, but is not harmed by 37°C (Daugherty et al., 1994). Testing *COR* gene expression after the 15°C drop from 37°C to 22°C should show whether or not *COR* gene induction is responsive to a change in temperature, especially since the 14°C drop from 22°C to 8°C was enough to induce a high level of expression of all of the *COR* gene families tested.

The finding that message first showed an increase over background at 16°C was unexpected. Initial work had only shown induction at 10°C or 12°C and colder, and had shown that the *COR*15 gene family was induced at a higher temperature than the other families (Thomashow et al., 1990). However, the plants used in those experiments had been grown in petri plates and not in pots (Horvath, personal communication), which may account for the difference. There are differences in humidity, temperature transfer and nutrient availability, to name a few, between the two growth methods.

The induction of *COR* genes at temperatures as high as 16°C also recalls the question of whether they are involved in acclimation to freezing or in the adjustment to metabolism at low, nonfreezing temperatures. Perhaps they allow physiological changes that are beneficial for both. Cold acclimation is usually effected at 2 or 4°C, although wheat has been shown to acclimate somewhat under a 10°C day/8°C night temperature regime (Gusta et al., 1982), and potato has shown increased freezing tolerance after exposure to temperatures as high as 12°C (Chen and Li, 1980). In each case, the level of acclimation achieved at these higher temperatures was inferior to that reached when lower temperatures were used, but was still substantial. It is not known at what temperatures above 4°C *Arabidopsis* will acclimate. This could be tested, although it seems likely that the temperatures of *COR* gene induction and of acclimation would overlap enough to leave the question of *COR* gene function open. More questions regarding the function of these *COR* genes could be probably be answered using plants that fail to express all members of any of the *COR* gene families. Antisense techniques have not yet been successful in producing such plants.

The temperature induction profiles described here should be useful in designing future experiments in which the standard temperatures used for low temperature induction (2-4°C) prohibitively slow plant growth and development or in which a weaker induction is preferred.

These profiles may also give direction to experiments intended to elucidate the signal operative in *COR* gene induction. Whatever the signal is, it must be responsive to temperature as such, be responsive to a temperature as high as 16°C, and be differentially responsive to different temperatures. Perhaps the signals at different temperatures are similar, but comprised of different components. Also, it would appear likely that the same signal is responsible for inducing all four *COR* gene families, since they all show the same pattern of induction. This is consistent with the observation that the core (CCGAC) of a promoter element shown to be sufficient for low temperature and drought inducibility, the DRE (Yamaguchi-Shinozaki and Shinozaki, 1994), is found in all *COR* gene promoters examined to date (Thomashow, 1994, Wang et al., 1995).

Materials and Methods

Plant growth and stress treatment

Arabidopsis thaliana (L.) Heyn. was grown in controlled environment chambers at 22°C with a 24 hour photoperiod (about 120 μmol m^{-2} s^{-1}) as previously described (Gilmour et al., 1988). Humidity was not controlled. After about three weeks of growth, plants were exposed to lowered temperatures.

In the first experiment, the temperature setting of the chamber was changed to 16°C, then to 14°C, 12°C, 10°C, 8°C and 6°C, for 24 hours each. At the end of each 24 hour period, one pot of each genotype was removed and the plants harvested

and frozen in liquid nitrogen. This harvesting was complete within 15 minutes or less of removal from the chamber. The chamber stabilized at each new temperature within about 10 minutes of being reset (empirical observation). A Tempscribe® chart recorder placed in the chamber confirmed that the temperature changed quickly and displayed little or no fluctuation during each 24 hour period.

In the second experiment, one set (9 pots) of 1,6 plants was moved to a second chamber and maintained at 22°C, while the rest of the plants were left in the first chamber. After 24 hours, a pot of plants was harvested from each genotype in each chamber, the temperature setting in the first chamber was changed to 19°C, and one pot of 1,6 from the second chamber was moved back into the first chamber. After another 24 hours, the plants in the pot that had been moved into the 19°C chamber and one pot of each of the genotypes of plants that had been in the chamber were harvested. If there were too few pots for the number of samples needed, a half pot of plants was harvested. Another pot from the second chamber was moved into the first and the temperature was changed to 16°C. This sequence was repeated for 14°C, 12°C, 10°C, 8°C, 6°C and 4°C. Before the last pot of plants was moved into the 4°C chamber, half of the plants were harvested so that the long-term effect of having moved the plants into the second chamber could be monitored. The light levels in the second chamber were

slightly higher than those in the first and the plants that had been moved gradually became anthocyanic.

RNA extraction and fractionation

Total RNA was extracted essentially as described (Gilmour et al., 1988) with a few modifications. Frozen pulverized tissue was extracted with equal volumes of phenol/chloroform/isopropyl alcohol (25/24/1, v/v/v) and extraction buffer (1% w/v triisopropylnaphthelene sulfonic acid, 6% w/v p-aminosalicylic acid, 100 μM Tris-HCl pH 7.6, 50 mM EGTA pH 8, 125 mM NaCl, 1% SDS, 10 mM DTT) on ice. This mixture was further homogenized in the tube using a Tekmar® Tissumizer, centrifuged (10,000 rpm in the SA600 rotor), and the supernatant extracted again with phenol/chloroform/isopropyl alcohol. This second set of tubes and all subsequent tubes and solutions except ethanol were made free of RNase by treatment with DEPC (Sambrook et al., 1989). Nucleic acids were precipitated with cold 95% ethanol, resuspended in 1 ml double distilled water, transferred to a microfuge tube and precipitated on ice for an hour with 1/4 volume 10 M LiCl. The pellets were precipitated again with 95% EtOH and resuspended in 200 μl double distilled water. The OD_{260} was measured to estimate the concentration of RNA.

For fractionation, RNA (5 to 40 μg) was dried down and resuspended in formaldehyde loading buffer containing EtBr (about 100 ng/ml), incubated at

68°C for 10-15 minutes and fractionated on denaturing formaldehyde agarose gels (Sambrook et al., 1989). The RNA was then transferred (Sambrook et al., 1989) to Magna NT membranes (MSI) using 10x SSPE made according to instructions provided by Schleicher and Schuell. This recipe for SSPE is slightly different from that described by Sambrook et al. (1989). RNA was UV cross-linked (Stratalinker, Stratagene) to the filters.

For northern hybridizations, blots were prehybridized, hybridized and washed in a Robbins Scientific Model 400 hybridization oven according to standard methods (Ausubel et al., 1987) with some modifications. Blots were incubated with prehybridization solution (50% formamide, 5x SSPE, 50 mM Potassium phosphate pH 8, 5x Denhardt's solution, 0.5% SDS, 100 μg/ml sheared, denatured fish sperm DNA) at 42°C for three hours to overnight and in hybridization solution (50% formamide, 5x SSPE, 50 mM potassium phosphate pH 6.5, 1x Denhardt's solution, 0.5% SDS, 5% dextran sulfate, 100 μg/ml sheared denatured fish sperm DNA) at 42°C overnight. Blots were rinsed and then given two 30-minute washes at room temperature in 2x SSPE/0.5% SDS, then washed two to three times, for 15 minutes each time, with 0.1x SSPE/0.5% SDS at 50°C.

Probes were made from gel-isolated fragments labeled with ^{32}P by random priming (Feinberg and Vogelstein, 1983). An *Eco* RI fragment from pHH67 was used to visualize *COR*15 message, a *Sac* I/*Hind* III fragment from pBI101.3

(Jefferson, 1987) served as a GUS probe, and *Eco* RI fragments from the cDNA clones pHH7.2 (Gilmour et al., 1992), pHH28 (Hajela et al., 1990) and pHH29 (Gilmour et al., 1992) were used to detect message from *COR47*, *COR78*, and *COR6.6*, respectively.

Literature Cited

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current Protocols in Plant Molecular Biology. Greene Publishing Associates and Wiley Interscience, New York, Chichester, Brisbane, Toronto, Singapore

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. 24: 701-713

Browse J, Kunst L, Anderson S, Hugly S, Somerville C (1989) A mutant of *Arabidopsis* deficient in the chloroplast 16:1/18:1 desaturase. Plant Physiol. 90:522-529

Chen H-H, Li PH (1980) Characteristics of cold acclimation and deacclimation in tuber-bearing *Solanum* species. Plant Physiol. 65:1146-1148

Daughtery CJ, Rooney MF, Paul A-L, de Vetten N, Vega-Palas MA, Lu G, Gurley WB, Ferl RJ (1994) Environmental stress and gene regulation. In, EM Meyerowitz, CR Somerville, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 769-806

Ding JP, Pickard BG (1993) Modulation of mechanosensitive calcium-selective cation channels by temperature. The Plant Journ. 3:713-720

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6-13

Gusta LV, Fowler DB, Tyler NJ (1982) Factors influencing hardening and survival in winter wheat. In, PH Li, A Sakai, eds, Plant cold hardiness and freezing stress, mechanisms and crop implications, v. 2. Academic Press, New York, London, Paris, San Diego, San Francisco, San Paulo, Sydney, Tokyo, Toronto, pp 23-40

Hajela RK, Horvath DP, Gilmour SJ, Thomashow MF (1990) Molecular cloning and expression of *cor* (cold-regulated) genes in *Arabidopsis thaliana*. Plant Physiol. 93:1246-1252

Horvath DP, McLarney BK, Thomashow MF (1993) Regulation of *Arabidopsis thaliana* L. (Heyn) *cor78* in response to low temperature. Plant Physiol. 103:1047-1053

Gilmour SJ, Hajela RK, Thomashow MF (1988) Cold acclimation in *Arabidopsis thaliana*. Plant Physiol. 87:745-750

Gilmour SJ, Artus NN, Thomashow MF (1992) cDNA sequence analysis and expression of two cold-regulated genes of *Arabidopsis thaliana*. Plant Mol. Biol. 18:13-21

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol Rep. 5:387-405

Knight MR, Campbell AK, Smith SM, Trewavas AJ (1991) Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. Nature 352:524-526

Knight H, Trewavas AJ, Knight MR (1996) Cold calcium signaling in Arabidopsis involves two cellular pools and a change in calcium signature after acclimation. The Plant Cell 8:489-503

Krishna P, Sacco M, Cherutti JF, Hill S (1995) Cold-induced accumulation of hsp90 transcripts in *Brassica napus*. Plant Physiol. 107:915-923

Miquel M, Browse J (1992) *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis. The Jour. Biol. Chem. 267:1502-1509

Monroy AF, Dhindsa RS (1995) Low-temperature signal transduction: Induction of cold acclimation-specific genes of alfalfa by calcium at 25°C. The Plant Cell 7:321-331

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A laboratory manual*, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Steponkus PL (1984) Role of the plasma membrane in freezing injury and cold acclimation. Ann. Rev. Plant Physiol. 35:543-584

Thomashow MF (1994) *Arabidopsis* as a model for studying mechanisms of plant cold tolerance. In, EM Meyerowitz, CR Somerville, eds., *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 807-834

Thomashow MF, Gilmour SJ, Hajela R, Horvath D, Lin C, Guo W (1990) Studies on cold acclimation in *Arabidopsis thaliana*. In, AB Bennett, SD O'Neill, eds., *Horticultural Biotechnology*. Wiley-Liss, New York, pp 305-314

Wang H, Datla R, Georges F, Loewen M, Cutler A (1995) Promoters from *kin1* and *cor6.6*, two homologous *Arabidopsis thaliana* genes: transcriptional regulation and gene expression induced by low temperature, ABA, osmoticum and dehydration. Plant Mol. Biol. 28:605-617

Welin BV, Olson A, Palva ET (1995) Structure and organization of two closely related low-temperature-induced *dhn/lea/rab*-like genes in *Arabidopsis thaliana* L. Heynh. Plant Mol. Biol. 29:391-395

Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of the expression of a desiccation-responsive *rd29* gene of *Arabidopsis thaliana* and the analysis of its promoter in transgenic plants. Mol. Gen. Genet. 236:331-340

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. The Plant Cell 6:251-264

APPENDIX

APPENDIX

SEQUENCE ALIGNMENT OUTPUT FROM PHDsec AND PHDthreader

PHDsec: Abbreviations

```
MAXHOM ALIGNMENT HEADER:  ABBREVIATIONS FOR SUMMARY
ID      :  identifier of aligned (homologous) protein
STRID   :  PDB identifier (only for known structures)
PIDE    :  percentage of pairwise sequence identity
WSIM    :  percentage of weighted similarity
LALI    :  number of residues aligned
NGAP    :  number of insertions and deletions (indels)
LGAP    :  number of residues in all indels
LSEQ2   :  length of aligned sequence
ACCNUM  :  SwissProt accession number
NAME    :  one-line description of aligned protein
```

PHDsec: COR15am

```
MAXHOM ALIGNMENT HEADER:  SUMMARY
ID          STRID IDE WSIM LALI NGAP LGAP LEN2 ACCNUM NAME
bla1_bacce        38   24   68    2   11   306 P10424 BETA-LACTAMASE PRECURSOR,
h1g_strpu         34   38   74    0    0   217 P07796 HISTONE H1-GAMMA, LATE.
tola_ecoli        33   35   89    0    0   421 P19934 TOLA PROTEIN.
edc8_dauca        32   33   88    1    3   555 P20075 EMBRYONIC PROTEIN DC8 (CL
tpma_xenla        32   24   85    2   18   284 Q01173 TROPOMYSIN ALPHA CHAIN,
dnak_bacme        32   27   82    2   14   605 P05646 DNAK PROTEIN (HEAT SHOCK
vg24_bpm15        32   27   79    1    8   132 Q05231 GENE 24 PROTEIN (GP24).
tpma_rante        31   24   86    2   18   284 P13105 TROPOMYSIN ALPHA CHAIN,
cyli_bovin        31   28   84    2    9   667 P35662 CYCLIN I.
le29_goshi        31   30   81    1    7   302 P13940 LATE EMBRYOGENESIS ABUNDA
cyli_human        31   26   88    2   16   598 P35663 CYCLIN (FRAGMENT).
vinc_human        31   19   85    3   15  1065 P18206 VINCULIN.
tpma_brare        31   23   85    2   18   284 P13104 TROPOMYSIN ALPHA CHAIN,
subv_bacsu        30   23   82    2   20   806 P29141 MINOR EXTRACELLULAR PROTE
lmb1_human        30   23   82    3   10  1786 P07942 LAMININ BETA-1 CHAIN PREC
modu_drome        31   26   78    2    8   544 P13469 DNA-BINDNG PROTEIN MODUL
rsp3_chlre        31   32   78    1    7   516 P12759 RADIAL SPOKE PROTEIN 3.
tola_haein        32   32   73    1    1   372 P44678 TOLA PROTEIN.
drpf_crapl        32   31   73    1    4   201 P23283 DESSICATION-RELATED PROTE
dyhc_yeast        30   22   80    2    5  4000 P36022 DYNEIN HEAVY CHAIN, CYTOS
sp2b_bacsu        31   25   74    1    7   332 P37575 STAGE II SPORULATION PROT
```

[PHDsec COR15am continued]

```
MAXHOM ALIGNMENT:  IN MSF FORMAT [small letters mark an insertion]
                   1                                                  50
predict_h191       AKGDGNILDD LNEATKKASD FVTDKTKEAL ADGEKAKDYV VEKNSETADT
bla1_bacce         .......... LQQNSTKKLD EVITYTKEDL VDYSP....V TEKHVDTGMT
h1g_strpu          .......... .TKKTKKVKK PAAKKAKKPA AKKPAAKKPA AKKAKKPAKK
tola_ecoli         ADAKAKAEAD AKAAEEAAKK AAADAKKKAE AEAAKAAAEA QKKAEAAAAA
edc8_dauca         AEVSRENTDY AYDKGREGGD VAAQKAEEAK EKAKMAKDTT MGKAGEYKDY
tpma_xenla         .KGTEDELDK YSEALKDAQE KLELSDKKAT DAEGDVAS1r AQERLSTALQ
dnak_bacme         .......IDK NGIVNVRAKD LGTNkiKSST GLSDDEIDRM VKEAEENADa
vg24_bpm15         ......TIDA FREEVKKksD DVTVELKPLL KLGQKAREAV VEVFKEFADI
tpma_brare         MKGTEDELDK YSEALKDAQE KLELAEKKAT DAEADVAS1r AQERLATALQ
cyli_bovin         AKKDTESTdd SKDAKKGKKE SKKDKKKDAK KDAASDAESG DSKDAKKDSK
le29_goshi         ........DV KNAAKGKSSE MRQATTEKAR ELADSAKENa kEKVRDMADR
cyli_human         .KYTKYTKKD TKKNAKKSSD AESEDSKDAK KDSKKVKKNV kkKKDVKKDT
vinc_human         PEGEEQIRGA LAEARkkERD DILRSLGEIS ALTSKLADLR RQgdSPEARA
tpma_brare         .KATEDELDK YSEALKDAQE KLELAEKKAT DAEGDVAS1r AQERLATALQ
subv_bacsu         ....GNSLNN PDWATSTALD WAMSEGVVAV TSNGNSGPNG WTVGseYAVT
lmb1_human         ......LLEE AKRASKSATD vtADMVKEAL EEAEKAQ.VA AEKAIKQADe
modu_drome         EEAAGLIDDE AEEDEEYNSD DEEDDDDDEL EPGEVSK... ....SEGADE
rsp3_chlre         AKMEAELQGK ELEAVRRRPT FVLRELKPAV ASADAVE... ....AAAAEL
tola_haein         .......... .EEAKAKAAE IAAQKAKQEA EAKAKLEAEA KAKAVAEAKA
drpf_crapl         .......... ..ASQSQGRQ QVSENAEDAK KKFSETTDSL KHKTSEATDS
dyhc_yeast         .......VNE LNKTLSKKST ELTEKEKEAR S..TLDKMLM EQNESERKQE
sp2b_bacsu         ..GTGLGLFA LNISGNKEAS APASLEDSLG SQTAKAGD.. .....TSADK

                   51                                                 89
predict_h191       LGKEAEKAAA YVEEKGKEAA NKAAEFAEGK AGEAKDATK
bla1_bacce         LGEIAEAAVR YSD.......  NTAGNILFHK IGGPKGYEK
h1g_strpu          VAKPAKKAAA KPAKKAAKPA KKAAKPAKKA AKPAK....
tola_ecoli         LKKKAEAAEA AAAEARKKAA TEAAEKAKAE AEKKAAAEK
edc8_dauca         TAQKAEEAKE KAAQKAEETK EKAGEytAQK AGEAKDTT.
tpma_xenla         KLEEAEKAAD ESERGMKVIE NRamELQEIQ LKEAKH...
dnak_bacme         lRNEADQLVF TTEKTLKDLE GKVEEAEVTK ANEAKDALK
vg24_bpm15         PDLEEDDDDE LVDEYSLQVC DIIAKAFRLI ATKPK....
tpma_rante         KLEEAEKAAD ESERGMKVIE NRaiELQEIQ LKEAKH...
cyli_bovin         KGKKDSK...  ..KDNKKKDA KKDAESTDAE SGDSKDAKK
le29_goshi         TSEMTNEAQE RGARKAEEAK EVVAEKAEGA AEETKKKNE
cyli_human         ESTDAESGDS KDERKdkKDT KKYPESTDTE SGDAKDARN
vinc_human         LAKQVATALQ NLQTKTNRap AKAAVHLEGK IEQAQ....
tpma_brare         KLEEAEKAAD ESERGMKVIE NRamELQEIQ LKEAKH...
subv_bacsu         FGssAKVMGY NKEDDVKALN NKEVELVEAG IGEAKD...
lmb1_human         lLTSIESETA ASEETLFNAS QRISELERNV EELKRKAAQ
modu_drome         VDESDDDEEA PVEKPVSKKS EKasEKSEEN RGIPK....
rsp3_chlre         TAQAEEAANA KWEADKAEAA EKARAEAEAA AEEQK....
tola_haein         KAEAEAKAKA AAEAKAkdAE AKAATEAKRK ADQA.....
drpf_crapl         ASHKANGAAR ETNDKAKETY NAAS....GK AGELKDKTQ
dyhc_yeast         ATEEIKKILK VQEekRKEVV MKSIQDIEPT ILEAQRGVK
sp2b_bacsu         QTSGAEKQAA QTEGTYKTYA VQAGKFSNEK GAE......
```

PHDsec: COR15bm

```
MAXHOM ALIGNMENT HEADER:  SUMMARY
ID          STRID IDE WSIM LALI NGAP LGAP LEN2 ACCNUM NAME
rla0_metva     38  33   64    1   14  336 P15826 ACIDIC RIBOSOMAL PROTEIN
crtc_caeel     35  33   68    2    8  395 P27798 CALRETICULIN PRECURSOR.
clpl_lacla     33  21   79    3   43  763 Q06716 ATP-DEPENDENT PROTEASE AT
if2_bacst      33  35   76    2   11  741 P04766 INITIATION FACTOR IF-2.
moes_pig       32  36   78    2    2  576 P26042 MOESIN (MEMBRANE-ORGANIZI
mysc_chick     31  26   81    2   33 1102 P29616 MYOSIN HEAVY CHAIN,CARDI
hs71_leima     31  22   81    3   46  634 P12076 MITOCHONDRIAL HEAT SHOCK
tebb_oxyno     31  24   85    2   13  385 P16458 SUBUNIT).
moes_mouse     31  36   78    2    2  576 P26041 MOESIN (MEMBRANE-ORGANIZI
moes_human     31  35   78    2    2  576 P26038 MOESIN (MEMBRANE-ORGANIZI
htr2_natph     30  27   79    3   19  534 P42259 PROTEIN II) (MPP-II).
ynp9_caeel     30  22   90    3   22  690 P34562 HYPOTHETICAL 79.2 KD PROT

MAXHOM ALIGNMENT:  IN MSF FORMAT [small letters mark an insertion]
                1                                                  50
predict_h208 VKSDGNILDD LNEATKKASD FVTDKTKEAL ADGEKTKDYI VEKTIEANET
rla0_metva   .........N DAKAVSVESA FITEKTADal AGDEALDDDL KEQISSSAVV
crtc_caeel   .......ITDS VEEAEAHAAE .....TFDKL KTVEKEKK.. .EKADEETRK
clpl_lacla   .......VIDR ENEIQKPAQK FCRRRKKNPL LVGEsgKTAV VEGllEAGTQ
if2_bacst    ..SDEEIFDD VKEAAKPA.. ....KKKGAA KGKETKRTEA QQQEKKAFQA
moes_pig     .......LKQ IEEQTKKAQQ ELEEQTRRAL ALEQERKRAQ SEAEKLAKeq
mysc_chick   ........TK LDEMTRLMND LTTQKTKLQS ENGEFVRQLe eTKSKNALAH
hs71_leima   .....NVIRV VNEPTAAALA YGMDKTKDSL IadLALSDYI LEEFRKTSGI
tebb_oxyno   ..RHTALQAA INKTVKGDN. .LVDISKVAD AAGKKGKVDA GIVKASAseG
moes_mouse   .......LKQ IEEQTKKAQQ ELEEQTRRAL ELEQERKRAQ SEAEKLAKeq
moes_human   .......LKQ IEEQTKKAQQ ELEEQTRRAL ELEQERKRAQ SEAEKLAKeq
htr2_natph   .......VKA LAEETKAATe tVQDRTQTTV DDIRETSDQv vEDTVDALee
ynp9_caeel   LKSENEKLLA KNEefKKKSH FveDETRKAI EKLEKSKVTI TELEQQADQT

                51                                                 90
predict_h208 ATEEAKKALD YVTEKGKEAG NKAAEFVEGK AEEAKNATKS
rla0_metva   ATEEAPKAET KKEEKKEEAA PAA..... ..........
crtc_caeel   AEEEARKKAE EEKEAKKDDD EEEKEEEEGH DE........
clpl_lacla   YRGSFEENIK QLVEEVKAAG nlAGEEVKGL ADIIK.....
if2_bacst    AKKKGKGPA. ....KGKKQA APAAKQVPQP AKKEKELPK.
moes_pig     EAEEAKEALL KAsdQKKTQE QLALEMAELT ARISQ.....
mysc_chick   ALQAARHDCD LLREQYEEEQ EAKAElsKGN AEVAQWRTK.
hs71_leima   DLSKERMALQ RvaEKAKCEL SSAMefITAN ADGAQH....
tebb_oxyno   NTATLKIADI FVQEKGKDAL NKAADHTDGA KVKGGAKGK.
moes_mouse   EAEEAKEALL QAsdQKKTQE QLASEMAELT ARISQ.....
moes_human   EAEEAKEALL QAsdQKKTQE QLALEMAELT ARISQ.....
htr2_natph   RTNDGIQEIN QSTDAQADAA QKATTMVEDM AATSEQ....
ynp9_caeel   RQEHFKTVED LASSRDKAEr eKTLKVLKSE LTESEKAHTT
```

PHDthreader [TOPITS]

```
TOPITS ALIGNMENTS HEADER: PARAMETERS
smin = -1.00 : minimal value of alignment metric
smax = 2.00  : maximal value of alignment metric
go   = 2     : gap open penalty
ge   = 0.2   : gap elongation penalty
len1 = 89    : length of search sequence, i.e., your protein

TOPITS ALIGNMENTS HEADER: ABBREVIATIONS
RANK         : rank in alignment list, sorted according to z-score
EALI         : alignment score
LALI         : length of alignment
IDEL         : number of residues inserted
NDEL         : number of insertions
ZALI         : alignment zcore;  note: hits with z>3 more reliable
PIDE         : percentage of pairwise sequence identity
LEN2         : length of aligned protein structure
ID2          : PDB identifier of aligned structure
NAME2        : name of aligned protein structure

TOPITS ALIGNMENTS HEADER: ACCURACY
: Tested on 80 proteins, TOPITS found the
: correct remote homologue in about 30% of
: the cases, detection accuracy was higher
: for higher z-scores (ZALI):
ZALI>0       : 1st hit correct in 33% of cases
ZALI>3       : 1st hit correct in 50% of cases
ZALI>3.5     : 1st hit correct in 60% of cases
```

PHDthreader: COR15am

```
TOPITS ALIGNMENTS HEADER: SUMMARY
 RANK   EALI LALI IDEL NDEL   ZALI PIDE LEN2  ID2 NAME2
    1  81.08   89    5    1   2.37   22  153 1aep LIPOPHORIN III
    2  77.17   88   19    5   2.18   23  153 2gdm _ID: 1;
    3  76.58   88   19    5   2.15   23  153 2lh2 HEMOGLOBIN (AQUO,MET)
    4  76.58   88   19    5   2.15   23  153 1lh3 HEMOGLOBIN (CYANO,MET)
    5  75.35   87    3    2   2.09   15  144 1lpe LIPOPROTEIN-*E3 (/LDL$ RE
    6  75.08   88    4    2   2.08   15  421 1ses YL-TRNA SYNTHETASE (E.C.6
    7  75.07   88    2    1   2.07   15  139 1le4 LIPOPROTEIN-*E4 (/LDL$ RE
    8  74.68   88    4    2   2.06   15  421 1sry YL-TRNA SYNTHETASE (E.C.6
    9  74.55   89    9    3   2.05   12  146 1was TERIAL ASPARTATE RECEPTOR
   10  73.98   85    4    1   2.02   13  157 1bcf CTERIOFERRITIN (CYTOCHROM
   11  71.60   88   17    3   1.90   17  197 1col LICIN *A (C-TERMINAL DOMA
   12  71.55   89   14    3   1.90   17  289 1bab MOGLOBIN THIONVILLE ALPHA
   13  71.40   89    4    1   1.89   17  384 1htm MAGGLUTININ ECTODOMAIN (S
   14  71.40   88   35    3   1.89   16  170 1fha RRITIN (H-CHAIN) MUTANT (
   15  71.03   86   14    3   1.88   20  127 2ccy TOCHROME $C(PRIME)
   16  70.83   88   20    3   1.87   16  118 2mhr OHEMERYTHRIN
   17  70.63   83   12    3   1.86   16  147 2hbg MOGLOBIN (DEOXY)
   18  70.47   88   22    5   1.85   19  524 1ddt PHTHERIA TOXIN (DIMERIC)
   19  70.42   89   24    5   1.85   18  148 1osa LMODULIN
   20  70.18   89   20    4   1.83   18  154 2spo OGLOBIN (MET) MUTANT WITH
```

PHDthreader: COR15bm

```
TOPITS ALIGNMENT HEADER:  SUMMARY
RANK  EALI LALI IDEL NDEL ZALI PIDE LEN2 ID2  NAME2
  1  78.87   90   38    4 2.15   20  153 1aep LIPOPHORIN III
  2  77.55   89   11    2 2.09   18  421 1ses YL-TRNA SYNTHETASE (E.C.6
  3  77.22   89   11    3 2.07   21  154 2spo GLOBIN (MET) MUTANT WITH
  4  77.20   89   11    4 2.07   24  154 2mge GLOBIN (MET) MUTANT WITH
  5  76.42   89   11    4 2.04   24  154 1mgn MYOGLOBIN MUTANT WITH INI
  6  75.42   90   10    2 1.99   22  421 1sry YL-TRNA SYNTHETASE (E.C.6
  7  75.20   90   17    4 1.98   12  584 1dlc TA-ENDOTOXIN CYRIIIA (BT1
  8  74.57   90   16    2 1.95   17  198 1abm GANESE SUPEROXIDE DISMUTA
  9  74.52   89   17    3 1.95   11  144 1lpe LIPOPROTEIN-*E3 (/LDL$ RE
 10  74.23   90   28    4 1.93   14  170 1fha RRITIN (H-CHAIN) MUTANT (
 11  73.87   89   30    7 1.92   30  153 1mbn OGLOBIN (FERRIC IRON-M
 12  73.83   90   27    5 1.92   19  508 1cpc PHYCOCYANIN
 13  73.83   90   27    5 1.92   19  508 1cpc PHYCOCYANIN
 14  73.77   90   35    7 1.91   22  197 1col LICIN *A (C-TERMINAL DOMA
 15  73.58   89    9    2 1.90   11  139 1le4 OLIPOPROTEIN-*E4 (/LDL$ R
 16  72.95   90   27    5 1.87   19  508 1cpc PHYCOCYANIN
 17  72.95   90   27    5 1.87   19  508 1cpc PHYCOCYANIN
 18  72.42   90    6    3 1.85   19  725 1vsg RIANT SURFACE GLYCOPROTEI
 19  72.30   88    8    1 1.84   11  384 1htm MAGGLUTININ ECTODOMAIN (S
 20  72.18   90   37    6 1.84   24  377 1dsb BA (DISULFIDE BOND FORMAT
```