LETTER TO THE EDITOR

GRP2 Proteins Contain Both CCHC Zinc Fingers and a Cold Shock Domain

As the number of reported sequences for gene products has grown, several protein domains have been identified that are associated with sequence-specific nucleic acid binding. Among these is the retroviral-type CCHC zinc finger, which is present twice in the human immunodeficiency virus (HIV) nucleic acid binding nucleocapsid protein (for review, see Coleman, 1992) and is included in a variety of eukaryotic proteins postulated to bind single-stranded nucleic acids (for examples, see Rajavashisth et al., 1989; Roussell and Bennet, 1993; Webb and McMaster, 1993). While searching sequence databases for proteins containing multiple CCHC zinc finger domains, we noted the occurrence of two of these zinc fingers in the similar glycine-rich proteins GRP2 of Arabidopsis (de Oliveira et al., 1990) and GRP2 of *Nicotiana sylvestris* (Obokata et al., 1991). In Figure 1A, the CCHC zinc fingers for the GRP2 proteins are aligned with the seven zinc fingers of the human cellular nucleic acid binding protein (Rajavashisth et al., 1989) and the two zinc fingers of the HIV retrovirus gag polyprotein. The GRP2 proteins share the conserved cysteine and histidine spacing as well as the glycines in positions 5 and 8 that are commonly found in proteins that contain more than one CCHC zinc finger.

Another nucleic acid binding domain, identified in bacteria as constituting the major cold shock–induced protein (Goldstein et al., 1990), is the "cold shock" domain. Several mammalian transcription factors contain regions similar to this domain (Wistow, 1990). Like the CCHC zinc fingers,

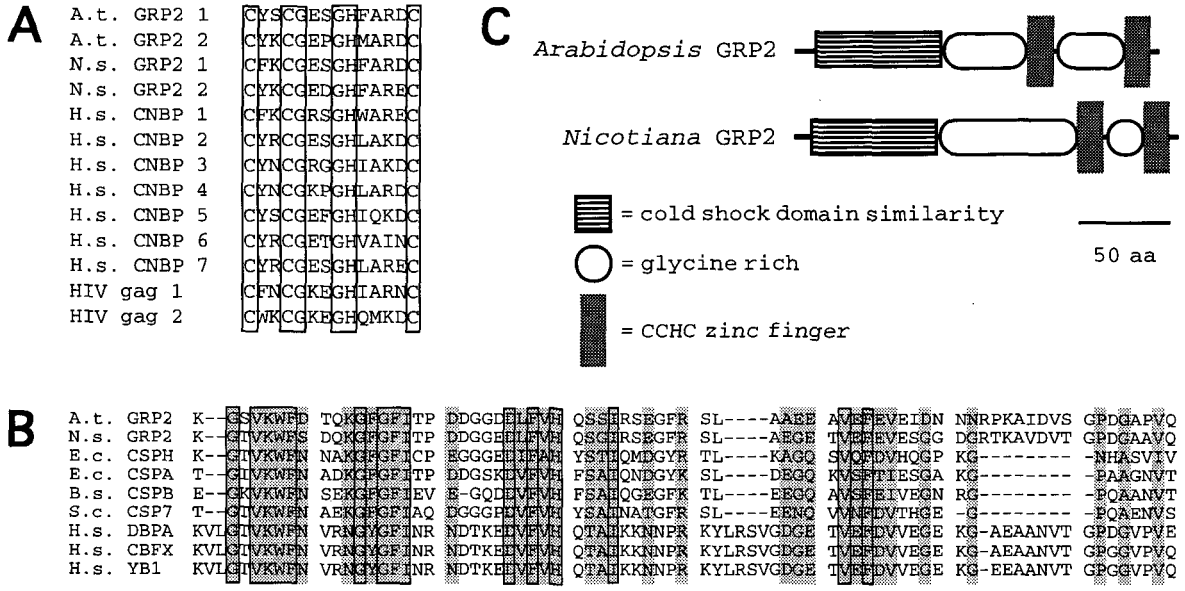

Figure 1. Structural Features Predicted for the GRP2 Proteins of Arabidopsis and *N. sylvestris*.

(A) Alignment of the two retroviral-type CCHC zinc fingers from the GRP2 proteins, the two zinc fingers from HIV gag polyprotein, and the seven human cellular nucleic acid binding protein (CNBP) zinc fingers. Identical amino acids are boxed.

(B) Alignment of the cold shock domains from the two GRP2 proteins and four bacterial and three human cold shock proteins (CSPs). Identical amino acids are boxed, and similar regions are shaded.

(C) Predicted structural features of the two GRP2 proteins depicted to scale diagramatically.

Accession numbers for the sequences shown in **(A)** and **(B)** are as follows: Arabidopsis GRP2, S47408; *N. sylvestris* GRP2, P27484; *Escherichia coli* CSPH, P24245; *E. coli* CSPA, P15277; *B. subtilis* CSPB, P32081; *Streptomyces clavuligerus* CSP7, Q01761; human DNA binding protein A (DBPA), P16989; human CCAAT box transcription factor (CBFX), P16990; human Y-box transcription factor (YB1), P16991; human CNBP, P20694; HIV gag polyprotein, P03348. All of the sequences are in the SWISS-PROT data base, with the exception of GRP2, which is in the GenBank data base.

LETTER TO THE EDITOR

the cold shock domain resembles regions of single-stranded nucleic acid binding proteins and can bind ssDNA in gel retardation studies (Schindelin et al., 1993; Schnuchel et al., 1993). Analysis of the two GRP2 proteins indicated that they also contain a conserved sequence similar to the DNA binding cold shock domain. The cold shock domains from four bacterial and three human proteins as well as amino acids 12–83 of Arabidopsis GRP2 and 10–81 of *N. sylvestris* GRP2 are aligned in Figure 1B. These domains share 15 identical and an additional 21 similar amino acids over the 62- to 77-amino acid structure. The only notable difference between the GRP2 proteins and the rest of the members is at position 10 of the alignment, where the GRP2 proteins have either an aspartic acid or a serine and the rest of the cold shock domain proteins have an asparagine. The three-dimensional structure that is available for one member of this group, *Bacillus subtilis* cold shock protein B (Schindelin et al., 1993; Schnuchel et al., 1993), indicates that this position lies at the border of the first β-strand and a turn of a five-stranded β-barrel. Thus, the lack of conservation at position 10 may not critically change the structure of this domain in the GRP2 protein as compared with other cold shock domains.

Two other plant proteins have been identified that have glycine-rich domains and whose mRNA levels increase in response to cold treatment (AtGRP7 and AtGRP8, van Nocker and Vierstra, 1993; also known as CCR1 and CCR2, Carpenter et al., 1994). Both proteins contain an RNA binding domain known as the RNP motif (or RRM) (Burd and Dreyfuss, 1994). However, neither contains a zinc finger motif or a cold shock domain.

The Arabidopsis and *N. sylvestris* GRP2 proteins are depicted schematically in Figure 1C. Over 90% of each protein is composed of the cold shock domain, the two CCHC zinc fingers, and the glycine-rich regions. The glycine-rich regions are reminiscent of similar "RGG box" sequences that are often found in combination with other nucleic acid binding domains in RNA binding proteins (Burd and Dreyfuss, 1994). Thus, it is likely that the GRP2 protein functions through recognition of specific nucleic acid sequences. To our knowledge, GRP2 is the first example of a protein that combines the well-defined nucleic acid binding motifs of both a cold shock domain and CCHC zinc fingers. The nucleic acid targets, developmental and inducible expression patterns, and function(s) of GRP2 remain to be elucidated.

Paul D. Kingsley
James Palis

University of Rochester
Department of Pediatrics and
Cancer Center
Rochester, NY 14642

REFERENCES

Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science **265,** 615–621.

Carpenter, C.D., Kreps, J.A., and Simon, A.E. (1994). Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. Plant Physiol. **104,** 1015–1025.

Coleman, J.E. (1992). Zinc proteins: Enzymes, storage proteins, transcription factors, and replication proteins. Annu. Rev. Biochem. **61,** 897–946.

de Oliveira, D.E., Seurinck, J., Inzé, D., van Montagu, M., and Botterman, J. (1990). Differential expression of five *Arabidopsis* genes encoding glycine-rich proteins. Plant Cell **2,** 427–436.

Goldstein, J., Pollitt, N.S., and Inouye, M. (1990). Major cold shock protein of *Escherichia coli.* Proc. Natl. Acad. Sci. USA **87,** 283–287.

Obokata, J., Ohme, M., and Hayashida, N. (1991). Nucleotide sequence of a cDNA clone encoding a putative glycine-rich protein of 19.7 kDa in *Nicotiana sylvestris.* Plant Mol. Biol. **17,** 953–955.

Rajavashisth, T.B., Taylor, A.K., Andalibi, A., Svenson, K.L., and Lusis, A.J. (1989). Identification of a zinc finger protein that binds to the sterol regulatory element. Science **245,** 640–643.

Roussell, D.L., and Bennet, K.L. (1993). glh-1, a germ-line putative RNA helicase from *Caenorhabditis*, has four zinc fingers. Proc. Natl. Acad. Sci. USA **90,** 9300–9304.

Schindelin, H., Marahiel, M.A., and Heinemann, U. (1993). Universal nucleic acid–binding domain revealed by the crystal structure of the *B. subtilis* major cold-shock protein. Nature **364,** 164–168.

Schnuchel, A., Wiltscheck, R., Czisch, M., Herrler, M., Willimsky, G., Graumann, P., Marahiel, M.A., and Holak, T.A. (1993). Structure in solution of the major cold-shock protein from *Bacillus subtilis.* Nature **364,** 169–171.

van Nocker, S., and Vierstra, R.D. (1993). Two cDNAs from *Arabidopsis thaliana* encode putative RNA binding proteins containing glycine-rich domains. Plant Mol. Biol. **21,** 695–699.

Webb, J.R., and McMaster, W.R. (1993). Molecular cloning and expression of a *Leishmania major* gene encoding a single-stranded DNA-binding protein containing nine "CCHC" zinc finger motifs. J. Biol. Chem. **268,** 13994–14002.

Wistow, G. (1990). Cold shock and DNA binding. Nature **344,** 823–824.