A Retrovirus-like Zinc Domain Is Essential for Translational Repression of Bacteriophage T4 Gene 32*

(Received for publication, November 26, 1990)

Youssef Shamoou2, Kevin R. Webster3, Kenneth R. Williams2, and William H. Konigsberg
From the Department of Molecular Biophysics and Biochemistry and the Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510

Gene 32 protein (gp32), a single-stranded DNA-binding protein from bacteriophage T4, contains a zinc-binding subdomain with sequence homologies to the 3-cysteine/1-histidine zinc-binding motif found in a variety of retroviruses and plant viruses. In vitro studies suggest that autoregulation of gp32 occurs at the level of translation by gp32 specifically binding gene 32 mRNA at an unusual stem-loop structure that can be modeled as an RNA pseudoknot. Nucleation of gp32 binding via this pseudoknot is thought to be needed to facilitate cooperative binding of gp32 through a largely unstructured region that overlaps the ribosome binding site (McPheeters, D. S., Stormo, G. D., and Gold, L. (1988) J. Mol. Biol. 291, 517-535). Removal of Zn(I1) from gp32 results in a protein that retains the ability to bind single-stranded RNA with high affinity but is unable to specifically autoregulate itself at the level of translation. Deletion of the pseudoknot sequences from the gene 32 autoregulatory region results in an mRNA that cannot be repressed by gp32. These results suggest that the zinc-binding subdomain of gp32 plays an essential role in autoregulation by providing a critical element necessary for nucleating cooperative binding at the gene 32 mRNA pseudoknot.

In recent years, a large number of nucleic acid-binding proteins have been identified that contain Zn(II), with at least two general classes of zinc proteins emerging. Perhaps the best characterized are the "zinc finger" proteins involved in sequence-specific double-stranded DNA binding using the (Tyr/Phe)-X-Cys-X2-X2-Phe-X2-Leu-X2-His-X2-His-X2 motif described by Berg (1) and for which a structure has recently been solved by Parraga et al. (2). A second class of zinc proteins use 3 cysteines and 1 histidine residue(s) as ligands (referred to as the CCHC box) (1). Examples of this class include a large number of nucleocapsid proteins from a variety of sources including the retroviruses human immuno-

nodeficiency virus (HIV) and murine leukemia virus (MuLV), as well as retrotransposons such as Drosophila copia and intracisternal A particles, and plant viruses tobacco rattle virus (TRV), pea early browning virus (PEBV), and cauliflower mosaic virus (Table I) (3-6). Bacteriophage T4 single-stranded DNA (ssDNA)-binding protein (gp32) also contains a 3-cysteine/1-histidine zinc-binding domain (Cys-X2-Cys-X2-Cys) (7). Characterization of the zinc function in gp32 has shown that it is important for stabilizing a specific domain within gp32 which, along with the N-terminal domain, contributes substantially to cooperative gp32-gp32 interactions (7, 8). This study also suggested that, within the physiological salt range, chemical removal of zinc from gp32 to produce apo-gp32 results in a protein that still binds ssRNA tightly but with reduced cooperativity (7, 8). Concurrent with our studies on the function of zinc in gp32, McPheeters et al. (19) began characterizing the structure of gene 32 mRNA. It has been previously shown that gp32 autoregulates itself at the level of translation and that gene 32 overexpression can only be achieved after the deletion of the 5' sequences of gene 32 mRNA that provide the target for repression (10, 11). von Hippel et al. (12) proposed that gp32 autoregulation is dependent on a "uniquely single-stranded" or AU-rich region at the 5' end of gene 32 mRNA. It has been recently proposed that an unusual stem-loop structure 5' to this AU-rich region acts as a nucleation point for gp32 binding to gene 32 mRNA (9). Phylogenetic comparisons of other T-even phage gene 32 mRNA sequences allowed McPheeters et al. (9) to propose that the 5' end of the gene 32 mRNA form a conserved pseudoknot structure. In this report, we present work that indicates that the zinc-binding subdomain of gp32 and a putative RNA pseudoknot in gene 32 mRNA are essential to gene 32 autoregulation. We suggest that the CCHC zinc-binding motif found both in T4 gp32 and in many nucleocapsid proteins may operate through a common mechanism of binding RNA pseudoknot structures.

EXPERIMENTAL PROCEDURES

Plasmids—Construction of plasmids pY55, g32(0), and g32(2wt) have been previously described (11). Plasmid g32(-38) was prepared for this work using a pair of complementary oligonucleotides that reconstructed the entire 5' region of g32 from position -38 to +3 and contained 5' EcoRI and 3' BamHI overhangs to facilitate cloning operations. Oligodeoxynucleotides were synthesized by the Yale Medical School Protein and Nucleic Acid Chemistry Facility. The oligonucleotides were phosphorylated and purified before being cloned into the EcoRI-BamHI sites of the commercial vector pBluescript International Biotechnologies, Inc. and sequenced.

Protein Purification and Quantitation—Expression of g32 and purification of native gp32 has been previously described and was quantitated using amino acid analysis (11, 13). Apo-gp32 was prepared essentially as described (7, 8). The apo-gp32 used in these experiments contained about 0.2 eq of Zn(II), as determined by atomic absorption and amino acid analysis. gp32 and apo-gp32 were dialyzed into 10 mM Tris, pH 8.0, 50 mM NaCl, 1 mM dithiothreitol, and 1 mM EDTA prior to use in the in vitro transcription/translation assays.

In Vitro Transcription/Translation Assays—DNA-dependent coupled transcription/translation reactions were performed using a kit from Squibb Company, Wallingford, CT 06492.

* The abbreviations used are: HIV, human immunodeficiency virus; MuLV, murine leukemia virus; TRV, tobacco rattle virus; PEBV, pea early browning virus; ss, single-stranded, wt, wild type.
Comparison of 3-cysteine/1-histidine zinc subdomains of T4 gp32 with viral nucleocapsid proteins

| Virus and potential RNA target* | Zinc subdomain sequence | Ref. |
|---------------------------------|-------------------------|-----|
| T4 gp32 (g32 mRNA)              | C TSS                   | C PV | C 7 |
| Animal retroviruses             | C FR                    | C GYDA | C PV | C 7 |
| HTLV (tRNA<sup>B</sup>)         | C PR                    | C GYD | C PV | C 7 |
| HTLV II (tRNA<sup>P</sup>)      | C FR                    | C GYD | C PV | C 7 |
| MuLV (tRNA<sup>B</sup>)         | C PL                    | C GDPB | C PV | C 7 |
| BaEV (tRNA<sup>P</sup>)         | C PL                    | C GDPB | C PV | C 7 |
| BLV (tRNA<sup>U</sup>)          | C PL                    | C GDPB | C PV | C 7 |
| FeLV (tRNA<sup>R</sup>)         | C PL                    | C GDPB | C PV | C 7 |
| SSV (tRNA<sup>U</sup>)          | C PL                    | C GDPB | C PV | C 7 |
| HIV-1 (tRNA<sup>S</sup>)        | C PL                    | C GDPB | C PV | C 7 |
| EIAV (tRNA<sup>R</sup>)         | C PL                    | C GDPB | C PV | C 7 |
| Visc (tRNA<sup>S</sup>)         | C PL                    | C GDPB | C PV | C 7 |
| MMTV (tRNA<sup>A</sup>)         | C PL                    | C GDPB | C PV | C 7 |

* HTLV, human T-cell lymphotropic virus; BaEV, baboon endogenous virus; BLV, bovine leukemia virus; FeLV, feline leukemia virus; SSV, simian sarcoma virus; EIAV, equine infectious anemia virus; MMTV, murine mammary tumor virus; IAP, intracisternal A particles; CaMV, cauliflower mosaic virus; FMV, fowlpox mosaic virus; CERV, carnation etched ring virus.

Results and Discussion

Using an in vitro coupled transcription-translation system and the mRNAs described in Fig. 1, we have shown that gp32 specifically represses translation of mRNA from constructs carrying the entire 5' autoregulatory region of gene 32 mRNA, with 60% repression occurring at approximately 12 μM gp32 (see Fig. 2). Levels of translation of other nontarget mRNAs within the same assay, such as the mRNA for β-lactamase, are unaffected even at 25 μM gp32. In addition, construct g32(-38), which contains none of the 5' autoregulatory sequences but still produces an amber peptide of gp32 (at position 115), is also nonrepressible by addition of gp32 to the translation mixture on 17.5% sodium dodecyl sulfate-polyacrylamide gels; the gels were dried and autoradiography was performed for 12-16 h at -80 °C.

Comparison of three-dimensional structures of T4 gp32 and apo-T4 gp32 with viral nucleocapsid proteins

Table I

![Fig. 1. Summary of mRNAs produced in in vitro transcription/translation assays. a, g32(2wt) contains all sequences required for efficient autoregulation of gene 32. The sequences required for the gene 32 pseudoknot are indicated in Box A, and the AU-rich sequences in Box B. Nucleotides proposed to be involved in pseudoknot base pairing are underlined. The initiation codon begins at position 0. (b) g32(20) contains a deletion of all the 5' autoregulatory sequences. The ribosome binding site of g32(20) is the same as that used in the gene encoding E. coli SSB. Both g32(2wt) and g32(20) are derived from plasmid pUC19 containing the λ pL promoter (11), c, g32(38) is derived from plasmid pB120 and contains only the AU-rich and ribosome binding site of gene 32 fused to β-galactosidase.](image-url)
the relative affinities of apo-gp32 and gp32 for double-stranded versus single-stranded RNA are about equal. This conclusion is consistent with the previous Nadler et al. (8) study, which predicted from salt sensitivity studies that below a salt concentration of about 0.15 M, apo-gp32 should bind as well or better than native gp32 to the ssRNA polyriboethanoxyadenlic acid. As seen in Fig. 3, apo-gp32 is unable to repress translation of gp32 amber 115 peptide from gene 32 mRNA containing the entire 5' autoregulatory region. There is about 15% repression observable at 25 μM apo-gp32 that we believe results from the 20% zinc contamination of our apo-gp32 preparations or from apo-gp32 itself being able to repress at these higher concentrations. Densitometry of the gel in Fig. 3 indicates that 60 μM apo-gp32 should result in 50% repression, 5-fold more than is required for zinc-containing gp32. In fact, 60 μM represents the level at which gp32 might be expected to bind nonspecifically to other RNAs. If apo-gp32 is reconstituted with Zn(II) and then dialyzed versus EDTA to remove excess Zn(II), translational repression is restored (Fig. 3). A partial proteolysis product of gp32 corresponds to residues 17–253 that binds tightly but noncooperatively to single-stranded nucleic acids and that will be referred to as "core gp32" was also tested for its ability to regulate gene 32 mRNA. The results from our experiments on core gp32 demonstrate that translation products are not measurably repressed even at core gp32 concentrations as high as 222 μM (data not shown). These results, taken together with our characterization of the gene 32 mRNA, suggest that apo-gp32 is impaired in its ability to specifically recognize and cooperatively bind to mRNA containing autoregulatory sequences that are sufficient for autoregulation by native gp32. It appears, therefore, that both cooperativity and specific RNA recognition are essential for autoregulation by gp32. Based on the NMR solution structure of HIV p55, which contains a CCHC zinc-binding subdomain similar to gp32 by Summers et al. (15) and microcalorimetry studies of gp32 and apo-gp32 by Keating et al. (14), it seems plausible that the Zn(II) ion may stabilize a subdomain within gp32 that is able to recognize the 5' double-stranded RNA structure in gene 32 mRNA through sets of interactions from the turns between the zinc ligands and "knuckles" of the zinc finger. Alternatively, both Karpel (16) and Nadler et al. (8) have speculated that the relative inability of apo-gp32 to initiate cooperative binding on a two-site lattice as opposed to propagating cooperative binding on a longer lattice may have a profound effect on gp32 autoregulation. Additional studies will be required in order to differentiate between these two possibilities.

It is interesting to note that recent studies on retroviral nucleocapsid protein p10 of MuLV have shown that p10 is able to bind and stabilize double-stranded RNA sequences (tRNA), whereas mutagenesis studies have demonstrated that disruption of the p10 zinc domain results in an inability of the virus to encapsidate viral RNA (17). It has also been shown that plant viruses of the tobavirus family, such as TRV and PEBV, contain a pseudoknot structure at the 3' end of the viral genome thought to be involved in viral RNA replication and packaging (5, 6). The TRV and PEBV 3' pseudoknot structures have striking homologies to tRNA although they cannot be aminoacylated by aminoacyl tRNA synthetase as in the case of the toba-, tymo-, bromo-, cucu-mo-, and hordeiviral RNAs (19). In light of our results indicating the potential importance of the gp32 CCHC zinc-binding motif to pseudoknot recognition, it is possible that this motif in retroviral proteins such as MuLV p10, HIV p15, and the plant virus cauliflower mosaic virus ORF IV protein may have a similar role in the recognition of known pseudoknot domains in tRNAs (5, 20, 21). Furthermore, in the case of plant virus nucleocapsid proteins such as PEBV and TRV, the zinc domain may be responsible for the recognition of the 3' pseudoknot of the viral genome. If true, this would suggest that the CCHC zinc domain may be a general motif for specific double-stranded RNA pseudoknot recognition much in the same way that the (Tyr/Phe)-X-Cys-X4-Cys-X-Phe-X-Leu-X4-His-X4-His-X2e motif is used for sequence-specific double-stranded DNA binding.

REFERENCES
1. Berg, J. M. (1990) J. Biol. Chem. 265, 6513
2. Parraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., and Klevitt, R. E. (1988) Science 241, 635
3. Fuettner, J., and Hohn, T. (1987) Trends Biochem. Sci. 12, 92
4. Ono, M., Toh, H., Miyata, T., and Awaysa, T. (1986) J. Virol. 55, 387
5. MacFarlane, S. A., Taylor, S. C., King, D. I., Hughes, G., and Davies, J. W. (1989) Nucleic Acids Res. 17, 2245
6. van Belkum, A., Cornelissen, B., Linhorst, J. B., Puy, C., and

\[ \text{REFERENCES} \]
1. Berg, J. M. (1990) J. Biol. Chem. 265, 6513
2. Parraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., and Klevitt, R. E. (1988) Science 241, 635
3. Fuettner, J., and Hohn, T. (1987) Trends Biochem. Sci. 12, 92
4. Ono, M., Toh, H., Miyata, T., and Awaysa, T. (1986) J. Virol. 55, 387
5. MacFarlane, S. A., Taylor, S. C., King, D. I., Hughes, G., and Davies, J. W. (1989) Nucleic Acids Res. 17, 2245
6. van Belkum, A., Cornelissen, B., Linhorst, J. B., Puy, C., and
Translational Repression of Bacteriophage T4 Gene 32

Bosch, L. (1987) *Nucleic Acids Res.* **15**, 2837
7. Giedroc, D. P., Keating, K. M., Williams, K. R., and Coleman, J. E. (1987) *Biochemistry* **26**, 5251
8. Nadler, S. G., Roberts, W. J., Shamoo, Y., and Williams, K. R. (1990) *J. Biol. Chem.* **265**, 10393
9. McPhetres, D. S., Stormo, G. D., and Gold, L. (1988) *J. Mol. Biol.* **201**, 517
10. Russel, M., Gold, L., Morissett, H., and O'Farrell, P. Z. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8844
11. von Hippel, P. H., Kowalczykowski, S. C., Lonberg, N., Newport, J. W., Paul, L. S., Stormo, G., and Gold, L. (1982) *J. Mol. Biol.* **153**, 795
12. Shamoo, Y., Williams, K. R., and Konigsberg, W. H. (1988) *Proteins Struct. Funct. Genet.* **4**, 1
13. Summers, M. F., South, T. L., Kim, B., and Hare, D. R. (1990) *Biochemistry* **29**, 329
14. Karpel, R. L. (1990) in *The Biology of Nonspecific DNA-Protein Interactions* (Rezvanz, A., ed) p. 103, CRC Press, Inc., Boca Raton
15. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420
16. Hull, R., Sadler, J., and Longstaff, M. (1986) *EMBO J.* **5**, 3083
17. Haenni, A. L., and Chapeville, F. (1980) in *Transfer RNA: Biological Aspects* (Soll, D., Abelson, J. N., and Schimmel, P. R., eds) pp. 539–556, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Dumas, P., Moras, D., Florentz, C., Giege, R., Verlaan, P., Van Belkum, A., and Pleij, C. W. A. (1987) *J. Biomol. Struct. & Dyn.* **4**, 707
19. Covey, S. N. (1986) *Nucleic Acids Res.* **14**, 623
20. Russel, M., Gold, L., Morissett, H., and O'Farrell, P. Z. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8844
21. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420
22. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420
23. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420
24. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420
25. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420