Vol. 106, No. 2, 1982 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 31, 1982 Pages 265-269

Amino Terminal Sequence of the Bacteriophage T5-Coded Gene A2 Protein

Jay W. Fox, Amy Barish, Clifford E. Snyder, and Rolf Benzinger

Departments of Microbiology (JWF and AB) and Biology (CES and RB), University of Virginia, Charlottesville, VA 22908

Received April 9, 1982

Summary The first twenty-eight amino acid residues from the amino terminal of the T5 phage-coded gene A2 protein were determined. Some evidence is presented which suggests the existence of two forms of the protein; one in the cytoplasm which may have a signal sequence and another form present in the outer membrane. The amino terminal A2 protein sequence shows some sequence homology to the amino terminal region of the T4 phage-coded gene 32 protein. Finally it is important to note that residues ten through twenty-one of the A2 protein are amino acids with low ambiguity codons which should facilitate in the DNA sequencing of the A2 gene.

Introduction An interesting feature of bacteriophage T5 DNA entry into the bacterium is the two-step nature of the process. The first step involves the interaction of the host cell membrane and T5 DNA to terminate the DNA transfer when only 8% of the DNA has been injected (1,2). In the second step the pre-early T5-coded proteins are synthesized from the partially injected DNA. These proteins then serve to transfer the remaining 92% of the DNA into the cell (3). Two pre-early proteins, the products of genes A1 and A2, are required for this second-step transfer (3). To begin to understand the molecular basis of bacteriophage DNA transfer, we have purified one of these proteins, the gene A2 product (8). The purified protein has the following unusual combination of macromolecular interactions: it forms a noncovalent dimer (4), it binds non-cooperatively to both single- and double-stranded DNA (5); it binds to the protein porin in the outer membrane of uninfected cells (6); and removes the protein from infected cells (6). Since the A2 protein

0006-291X/82/100265-05$01.00/0

interacts with both membrane proteins and DNA it falls into a class of proteins which are predicted to be amphipathic (7). To gain a better understanding of the protein and to test this prediction, we decided that one approach would be to determine the amino acid sequence of the A2 protein. Also the amino acid sequence studies could be of significance when examining the DNA sequence.

Materials and Methods The gene A2 protein was purified as previously described (4) to greater than 99% homogeneity as judged from sodium dodecyl sulfate polyacrylamide gels.

Amino acid analyses following acid hydrolysis (5.7M HCl) for 24h at 110°C were performed on a Beckman 121 CL amino acid analyzer.

The A2 protein was subjected to digestion with BrCN in 70% formic at a 1:1 (w/w) ratio of protein to BrCN. The digestion was carried out at 25°C for 16h and terminated by the addition of distilled water followed by drying on a roto-evaporator.

The N-terminal BrCN peptide was isolated using a Beckman 421 high performance liquid chromatograph and a Waters Associates μBondapak C-18 column (7.8mm x 30cm).

The native A2 protein and N-terminal BrCN peptide were sequenced with a Beckman 890C Sequencer using the 0.1M Quadrol program (Beckman, 121178). In order to avoid excessive extractive losses from the cup, Polybrene (3mg) was utilized in all sequencing experiments. Polybrene was loaded into the spinning cup and subjected to three Edman degradation cycles prior to adding the peptide to be sequenced. For the sequence analysis run on the native A2 protein 70 nmoles were used. In the sequence analysis on the N-terminal BrCN 50 nmoles were loaded into the cup. The phenylthiohydantoin amino acids were identified by thin layer chromatography (8) and high performance liquid chromatography (9).

Results and Discussion The amino acid sequence shown in Figure 1 was derived by several lines of evidence: 1. direct sequencing from the amino terminal of

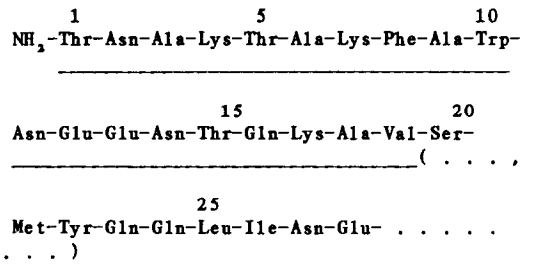

Fig. 1. Sequence Analysis of Amino Terminal Region of A2 Protein. Automatic Edman degradation using a Beckman 890C Sequencer with Polybrene and a 0.1M Quadrol program. The bar under the sequence represents the N-terminal BrCN peptide. Parenthesis represents residues not directly identified in this peptide.

the whole A2 protein. 2. BrCn cleavage of the protein followed by isolation of the amino terminal BrCN peptide and the amino acid analysis of the peptide (See Table 1) (which is also corroborated by the N-terminal sequence). 3. direct sequencing of the amino terminal BrCN fragment up to residue 19 out of a total of 21. At this point the remaining dipeptide from the BrCN peptide was washed out of the cup.

Several features of the amino terminal sequence are noteworthy. One is that the amino terminus residue is not methionine which is present in several cytoplasmic DNA-binding proteins encoded by coliphage and Escherichia coli. For example, the T4 phage coded gene 32 protein, fd and M13 phage coded single stranded DNA binding protein and the HU protein of E. coli all have methionine as the N-terminal residue in their protein sequences (10,11,12). The threonine N-terminal of the A2 protein suggests the possibility that there may exist a minor cytoplasmic form of the gene A2 protein which contains a signal sequence with an amino terminal methionine residue. This hypothesis is supported by the recent finding of the gene A2 protein in the outer membrane of T5 infected E. coli (Synder and Benzinger, unpublished). With reverse phase HPLC, we have observed that electrophoretically pure gene A2 protein is resolved into a major and minor peak. We are currently investigating whether the minor peak is A2 protein with a signal sequence still attached and the possibility that the two forms of the protein may have different functions.

The amino terminal sequence of gene A2 protein contains two residues, tryptophan 10 and methionine 21, which both possess only a single codon. Fur-

Table I. Amino Acid Composition of the BrCN N-Terminal Peptide from T5 Gene and A2 Protein

Amino Acid	BrCN Peptide
Asx	2.6[a] (3)[b]
Thr	2.5 (3)
Ser	0.8 (1)
Hse	0.6 (1)
Glx	2.8 (3)
Pro	0
Gly	0
Ala	3.8 (4)
Cys	0
Val	1.1 (1)
Met	0
Ile	0
Leu	0
Tyr	0
Phe	1.2 (1)
His	0
Lys	3.3 (3)
Arg	0
Trp	N.D.[c] (1)[d]
Total	21

[a] nmole amount

[b] Integer value of residue/molecule

[c] N.D.: not determined.

[d] Estimated

thermore, six of the ten residues between these two amino acids possess only two codons. This low level of codon ambiguity should facilitate locating this region of the gene in the T5 chromosome.

Finally, the gene A2 protein contains the very unusual dipeptide sequence glutamate-glutamate at positions 12 and 13. The sequence glutamate-aspartate is found at residues 26 and 27 of the T4 phage-coded gene 32 protein which can dimerize and has been demonstrated to bind DNA (12). Both the gene A2 protein and the T4 gene 32 protein have tryptophan and phenylalanine residues near these acidic dipeptides. The gene A2 protein has a tryptophan at position 10 and a phenylalanine at position 8 while the gene 32 protein has a tryptophan at position 31 and a phenylalanine at position 23 (12).

Investigations are presently being carried out in our laboratory to elucidate the total amino acid sequence of the gene A2 protein so that more

sophisticated sequence comparisons of A2 protein with other similar proteins can be made. We also are continuing to search for an A2 protein with a signal sequence attached so that we can further characterize the in vivo role these proteins play during infection.

References

1. Labedan, R,. (1978) Virology 85, 487-493.

2. Shaw, A.R., Lang, D., and McCorquodale, D.J. (1979) J. Virol. 29, 220-231.

3. McCorquodale, D.J. (1975) Crit. Rev. Microbiol. 4, 101-159.

4. Snyder, C.E. Jr., and Benzinger, R.H. (1981) J. Virol. 40, 248-257.

5. Snyder, C.E. Jr., Benzinger, R.H., and Burkhardt, H.J. (1980) Abstracts of the Cold Spring Harbor Bacteriophage Meetings, 29.

6. Snyder, C.E. Jr. (1982) Abstracts of the Ann. Meeting of the Am. Soc. for Microbiol., H11.

7. Kornberg, A. (1980) in "DNA Replication" Chapter 13. W.H. Freeman San Francisco.

8. Fox, J.W., Elzinga, M., and Tu, A.T. (1978) Biochemistry 18, 678-684.

9. Hunkapiller, M.W., and Hood, L.E. (1980) Science 207, 523-525.

10. Nakashima, Y., Dunker, A.K., Martin, D.A., and Konigsberg, W.H. (1974) FEBS Letters 43, 125-127.

11. Laine, B., Sautiere, P., Bisette, G., Cohen-Solal, M., Gros, F., and Rouviere-Yamiv, J. (1978) FEBS Letters 89, 116-120.

12. Williams, K.R., Lopresti, M.B., Setoguchi, M., and Konigsberg, W.H. (1980) Proc. Natl. Acad. Sci. USA 77, 4614-4617.