Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 6105–6109, October 1983
Microbiology

Efficient expression of influenza virus NS1 nonstructural proteins in *Escherichia coli*

(promoter/codon usage/expression vector/negative-strand RNA virus)

J. F. YOUNG*†, U. DESSELBERGER*‡, P. PALESE*, B. FERGUSON†§, A. R. SHATZMAN†§, AND M. ROSENBERG†§

*Department of Microbiology, Mount Sinai School of Medicine of the City University of New York, New York, NY 10029; §Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205

Communicated by Edwin D. Kilbourne, July 5, 1983

ABSTRACT **RNA segment 8 of the influenza A virus genome codes for two nonstructural proteins, NS1 and NS2, for which the functions are unknown. Cloned cDNA copies of this gene from three different influenza A virus strains were inserted into an *Escherichia coli* plasmid expression vector, pAS1, carrying the strong regulatable λ phage promoter, P_L. After induction, the NS1 proteins were overproduced to levels of 20–25% of total cellular protein. This was surprising in that the codon composition for these eukaryotic genes is similar to that for weakly expressed proteins in *E. coli*. Thus, under the appropriate conditions, it appears that high level expression of genes containing a relatively large proportion of minor codons can be obtained. The NS1 protein produced in bacteria from a cloned cDNA copy of the A/PR/8/34 virus *NS* gene was purified to apparent homogeneity and used to generate a high-titer monospecific rabbit antiserum. Immunoprecipitation studies showed this antibody to be crossreactive against the NS1 proteins produced by several different influenza A virus strains. Immunofluorescence experiments in Madin–Darby canine kidney cells showed the NS1 proteins to be located in the nucleoplasm early in infection for all strains examined. With some of the strains, NS1-specific immunofluorescence was observed predominantly in the nucleoli later in infection. This technology can be used to obtain other viral proteins in pure form for structural, functional, and immunological studies.**

Influenza A viruses contain a genome of eight segmented single-stranded RNAs of negative polarity that, when transcribed into mRNA, code for at least 10 polypeptides (1, 2). Three of these gene products (NS1, NS2, and M2) are found only in virus-infected cells. The other seven proteins are present in the virion, associated with the viral envelope (HA, NA, and M1), or as part of the core complex (PB1, PB2, PA, and NP). The ability to obtain large quantities of these viral proteins in pure form will greatly enhance studies of the structure of these polypeptides and help to elucidate the function(s) of these gene products during the virus replication cycle.

With the emergence of recombinant DNA technology, it has been possible to clone cDNA copies of the eight influenza virus genomic RNAs into *Escherichia coli* plasmid vectors (3). In this paper, we describe the bacterial production of the influenza virus NS1 nonstructural proteins from three different viral strains. The viral cDNA coding for the NS1 protein was inserted into a plasmid vehicle, pAS1, that contains transcriptional and translational signals derived from bacteriophage λ (4). Under the appropriate induction conditions, large quantities of NS1 protein were produced in an *E. coli* host. The protein was then extracted and purified for use as an immunogen to generate a high-titer monospecific antiserum in rabbits for *in vitro* and *in vivo* studies.

Immunoprecipitation studies using this antiserum showed it to be crossreactive with all strains tested, and immunofluorescence experiments revealed differences in the intranuclear location of NS1 depending on the strain of virus used to infect cells.

MATERIALS AND METHODS

Bacterial Strains and Plasmid Vectors. The cloning and sequence analysis of the cDNA copies of the *NS* genes from human influenza A virus strains A/PR/8/34, A/FW/1/50, and A/FM/1/47 has been described (5, 6). Chimeric pBR322 plasmids containing these inserts were propagated in *E. coli* strain C600. Plasmid pAS1 and its derivatives were grown in *E. coli* strains N99(cI+) and N5151(cIts857) (4). Restriction enzyme digestion, isolation of DNA fragments, ligation, transformation with plasmid DNAs, and restriction mapping were carried out according to published procedures (7).

Synthesis and Purification of NS1 Protein. The NS1 protein was purified from *E. coli* by following a modification of the procedure described for the cII protein (8). A 750-ml culture of *E. coli* strain N5151(CIts857) carrying plasmid pAS1/PR8 (see *Results*) was grown at 32°C in LB broth to an OD_{650} of 0.6. Two hundred and fifty milliliters of medium preheated to 65°C was then added with vigorous agitation, and incubation was continued at 42°C for an additional 2 hr. The bacteria were pelleted (20 min at 12,000 × *g*), suspended in phosphate-buffered saline, and repelleted (20 min at 12,000 × *g*). The bacterial pellet was frozen at −70°C, thawed, and resuspended in 20 ml of buffer A [50 mM Tris·HCl, pH 8/2 mM EDTA/0.1 mM dithiothreitol/5% (vol/vol) glycerol]. Lysozyme was added to a final concentration of 0.2 mg/ml, and the mixture was incubated on ice for 20 min. The suspension was centrifuged at 12,000 × *g* for 30 min, and the pellet was resuspended in 20 ml of buffer A. Sodium deoxycholate was added to a final concentration of 0.05%, and this mixture was homogenized in a Dounce homogenizer (15 strokes), incubated at 15°C for 30 min, and centrifuged at 12,000 × *g* for 45 min. The supernatant was collected and 5 M NaCl and 1 M $MgCl_2$ were slowly added to a final concentration of 1 M and 100 mM, respectively. After mixing for 1 hr at 4°C, the precipitate was collected by centrifugation (12,000 × *g*, 1 hr). The pellet was resuspended in 20 ml of buffer B (40 mM Tris·HCl, pH 7.5/2 mM $MgCl_2$). $MgCl_2$ was again slowly added to a final concentration of 100 mM, and the precipitate was col-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

† Present address: Molecular Genetics Division, Smith Kline & French Laboratories, Philadelphia, PA 19101.

‡ Present address: University of Glasgow, Institute of Virology, Glasgow, Scotland.

lected as described above. The pellet was dissolved in 5 ml of buffer B and dialyzed for 16 hr at 4°C in buffer B. By using this protocol, 5–10 mg of NS1 have been routinely obtained. Analysis of 25 μg of this material by polyacrylamide gel electrophoresis and Coomassie blue staining did not reveal any contaminating proteins (data not shown). Recently, large scale fermenter synthesis of the NS1 protein was achieved after induction at an OD_{650} of 4–8 (D. Wareheim, personal communication). This will permit the purification of 0.5–1.0 g of NS1 protein from a 10-liter culture.

Production of an Antiserum to the A/PR/8/34 NS1 Protein Synthesized in *E. coli*. Purified NS1 protein was injected into the medial ear vein of a rabbit in three doses of 100 μg each over a 5-week period. Blood was collected by cardiac puncture, followed by coagulation at 4°C overnight, and centrifugation at 5,000 × *g* for 30 min to remove the clotted cells. The serum was removed and used for immunoprecipitation and immunofluorescence assays.

RESULTS AND DISCUSSION

The *NS* gene of A/PR/8/34 (H1N1) influenza virus was previously cloned in pBR322 and the sequence was determined (5). This gene contains a 5′ leader sequence of 26 nucleotides, an open reading frame of 690 bases coding for the NS1 protein, and a 3′ untranslated region of 174 bases. It was of interest to express the NS1 protein in *E. coli* as a means of obtaining large quantities of this polypeptide free of contamination with other viral products and eukaryotic host components.

Plasmid expression vector, pAS1, has been developed for the efficient production of proteins in *E. coli* (4). This system has been used to synthesize large quantities of *E. coli lacZ*, simian virus 40 small tumor antigen, mammalian metallothioneins (9), and, more recently, the adenovirus *E1a* gene product (unpublished data). It derives signals from λ phage DNA to drive the transcription and translation of inserted foreign genes. The plasmid contains the λ promoter, P_L, an N utilization site (Nut) to relieve transcriptional polarity effects (when N protein is provided), the ribosome binding site from the *cII* gene, and the initiation codon for the cII protein. This initiation codon is adjacent to a unique *Bam*HI restriction endonuclease site. Cleavage of the plasmid with *Bam*HI and "in-phase" insertion of the influenza virus cDNA would permit expression of the viral protein using the cII translation signals.

The *NS* gene of A/PR/8/34 has a unique *Bam*HI cleavage site between the first and second codons (Fig. 1). Therefore, cleavage of plasmid pAPR801 with *Bam*HI liberates a 1,236-base-pair fragment that, when ligated into *Bam*HI-digested pAS1, results in the in-phase insertion of the gene into the vector regenerating the precise NS1 coding region (Fig. 1). An important feature of this system is that plasmids containing the P_L promoter must be introduced into a λ lysogenic host to stabilize the plasmid DNA (9). Cloning into lysogens also precludes the synthesis of proteins that might be toxic to the cells (10). If the lysogen used carries a temperature-sensitive mutation in the *cI* gene (11), P_L-directed transcription can be regulated. Bacteria are grown at 32°C to the appropriate density, and, after a shift-up of the culture to 42°C to inactivate the repressor, the synthesis of the desired gene product is turned on.

As shown in Fig. 2 (lanes 5–8), when pAS1/PR8 is introduced into an inducible lysogen, production of NS1 protein is observed, which accounts for ≈20% of the total cellular protein after only a 90-min induction period. Similar results were obtained when the *NS* genes from viruses A/FW/1/50 and A/FM/1/47 (6) were inserted and expressed in pAS1 (Fig. 2, lanes 1–4 and 9–12).

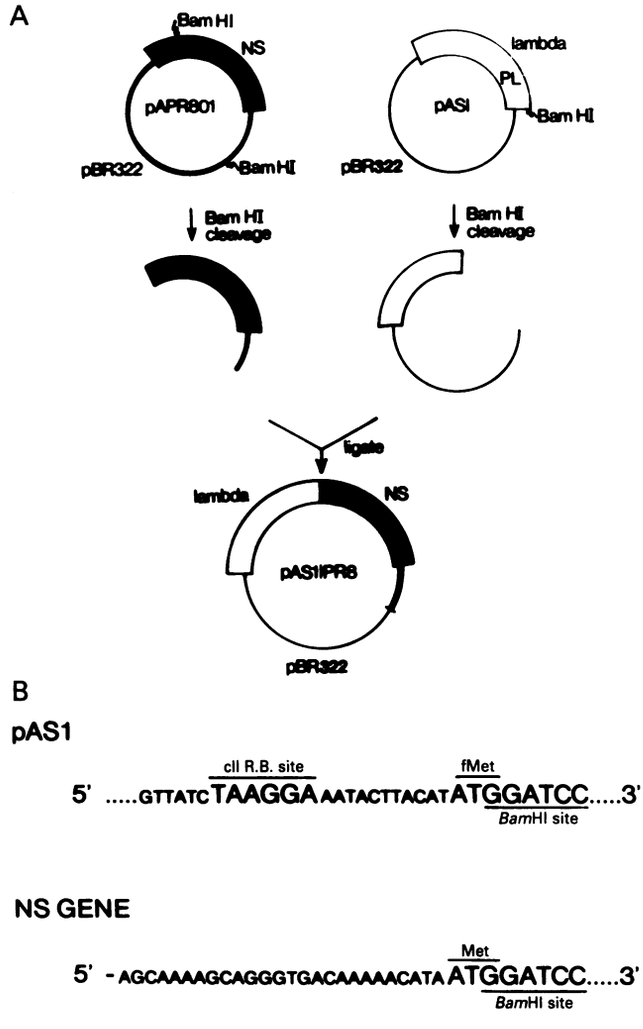

FIG. 1. Insertion of the influenza virus *NS* gene into pAS1. (*A*) Plasmid pAPR801 (5) was cut with restriction endonuclease *Bam*HI. The 1,236 base-pair fragment containing the *NS1* coding region was isolated by polyacrylamide gel electrophoresis and ligated into *Bam*HI-digested pAS1 DNA (4). This DNA was then used to transform *E. coli* λ lysogen N99CI+. (*B*) The sequences of the *cII* gene ribosome binding site (cII R.B. site) and the formylmethionine initiation codon (f-met) of pAS1 are indicated. The *NS* gene sequences located upstream to the unique *Bam*HI site are also shown. The coding sequences of the *NS* gene downstream to the *Bam*HI site were inserted into the unique *Bam*HI site of pAS1.

To date, the synthesis of proteins in *E. coli* has found limited application in influenza virus research. Three groups have reported the expression of influenza virus hemagglutinin molecules in *E. coli* (12–15). The hemagglutinin cDNA clones were inserted into a plasmid vehicle next to a prokaryotic promoter sequence and expressed in *E. coli*. The hemagglutinin proteins produced were hybrids containing amino acids encoded by the prokaryotic as well as by viral sequences. In contrast, because of the way the *NS1* coding region was inserted into the pAS1 vector, the NS1 protein produced in the bacteria has the same primary amino acid sequence as the protein produced during normal virus replication. This may be important when viral proteins produced in *E. coli* are examined for biological activity. Fig. 3 shows that identical mobilities were observed for the NS1 proteins made in *E. coli* and in virus-infected cells. The NS1 mobility differences observed among strains reflects variation in the lengths of the open reading frames in the *NS* genes of the three viruses. The lengths of the NS1 proteins from the

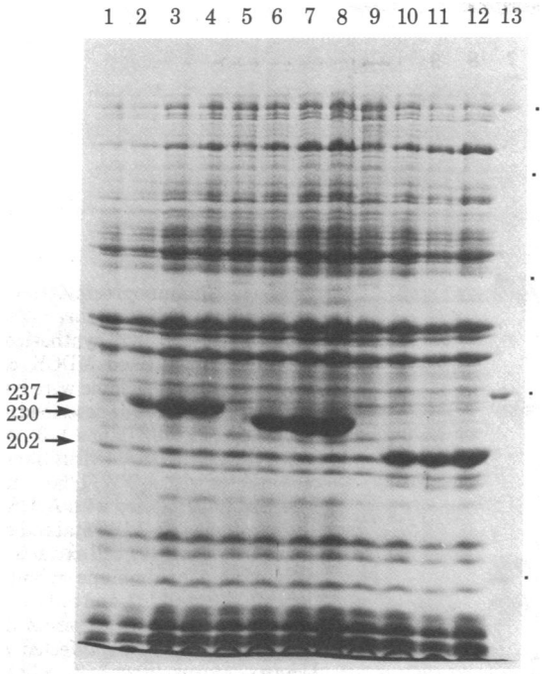

FIG. 2. Production of influenza virus NS1 proteins in *E. coli*. Bacteria transformed with plasmids containing the *NS* gene cDNA from strains A/FW/1/50 (lanes 1–4), A/PR/8/34 (lanes 5–8), and A/FM/1/47 (lanes 9–12) are shown. Cell extracts were prepared from 1-ml aliquots of culture, proteins were separated on a 10% polyacrylamide gel, and stained with Coomassie brilliant blue R-250. Uninduced cultures (lanes 1, 5, and 9), cultures induced for 30 min (lanes 2, 6, and 10), 60 min (lanes 3, 7, and 11), and 90 min (lanes 4, 8, and 12). Lane 13, Bio-Rad low molecular weight protein markers (10–100 kilodaltons). The predicted amino acid lengths of the three different NS1 proteins are indicated by arrows.

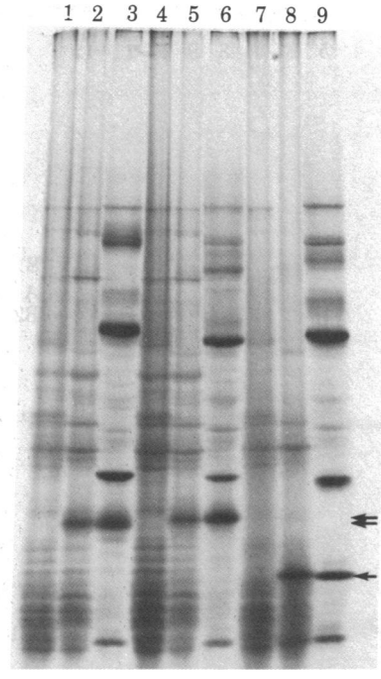

FIG. 3. Comparison of influenza virus NS1 proteins from infected cells with *E. coli*-produced NS1. Virus-infected cell proteins and *E. coli* proteins were labeled with [^{35}S]methionine as described (16–17). Cell extracts were prepared and polypeptides were separated on a 7–14% polyacrylamide gradient gel (16). Extracts of bacteria carrying plasmids containing the *NS* gene-specific cDNA either from influenza strain A/FW/1/50 (lanes 1 and 2), A/PR/8/34 (lanes 4 and 5), or A/FM/1/47 (lanes 7 and 8) are shown. Uninduced cultures (lanes 1, 4, and 7), induced cultures (lanes 2, 5, and 8). Extracts from MDCK cells infected with A/FW/1/50 (lane 3), A/PR/8/34 (lane 6), and A/FM/1/47 (lane 9) are shown for comparison. The arrows indicate the positions of the three different NS1 proteins.

three strains are 237, 230, and 202 amino acids for A/FW/1/50, A/PR/8/34, and A/FM/1/47, respectively (5, 6, 17).

Codon Usage in the *NS1* Gene. Recent studies on the evaluation of codon usage and tRNA abundance in *E. coli* have shown a correlation between expression levels of natural genes and their codon composition (18–22). It was observed that highly expressed genes use a large percentage of codons that are read by abundant tRNAs, and weakly expressed genes are translated by a higher proportion of minor tRNA species (18–22). It was, therefore, interesting to find that the *NS1* gene resembles a weakly expressed *E. coli* gene, being composed of >30% minor codons. The efficient expression of this gene suggests that rare codon usage may not be a limiting factor in obtaining high-level synthesis of a protein in *E. coli* under the conditions used in this system. Rather, other factors such as initiation of translation, messenger RNA stability, and protein stability may be operative, and it is only under certain growth conditions that codon composition may be important.

Immunoprecipitation of Viral NS1 Proteins with Antibody to *E. coli* Produced NS1. A rabbit antiserum was generated to purified NS1 protein made in *E. coli*. This antiserum was then used for immunoprecipitation studies of Madin–Darby canine kidney (MDCK) cells infected with various strains of influenza virus. Fig. 4 shows the autoradiograms of one such experiment. The sequences of the *NS* genes of the strains used have been analyzed (5, 6, 24, 25), and their predicted NS1 proteins differ by as little as 4.5% (A/FW/1/50) or as much as 33% (A/duck/Alberta/60/76). Quantitative precipitation was observed that was specific for the NS1 proteins from all of the strains examined, which shows crossreactivity. However, because competition experiments were not performed using the various strains, the degree of crossreactivity cannot be assessed. Previous studies using either mouse convalescent antiserum or antiserum against NS1-specific cytoplasmic inclusions have shown similar crossreactivities (26, 27).

Precipitation of the NS proteins from influenza B or influenza C virus-infected cells was not observed under these conditions (data not shown).

Intracellular Location of the NS1 Protein in Virus-Infected Cells. Previous studies had suggested that the NS1 protein is associated with the nucleolus (26–29). In those studies, however, NS1 protein from infected cells was an immunogen; therefore, those sera might contain antibody to other viral proteins or eukaryotic components. By using the monospecific antiserum made against the *E. coli*-produced NS1 protein, it was possible to reexamine this question. MDCK cells were infected with either A/Bel/42 or A/PR/8/34 virus, harvested at various times, and stained with anti-NS1 antibody. Two hours after infection both strains showed nucleoplasmic immunofluorescent staining for the NS1 protein (Fig. 5 *A* and *B*). This pattern of staining was also observed 4 hr and 6 hr after infection with A/PR/8/34 virus (Fig. 5 *C* and *E*) and continued throughout infection until cell death occurred (data not shown). However, in A/Bel/42 virus-infected cells, staining for NS1 was observed throughout the nucleus at 4 hr (Fig. 5*D*), and thereafter the nucleoli were predominantly stained (Fig. 5*F*). This suggests that (*i*) there are strain differences in the NS1 proteins that dictate the ultimate intracellular location of the NS1 protein or (*ii*) the strains differ in their replication scheme or their effects on the host cell (or both), which results in the differ-

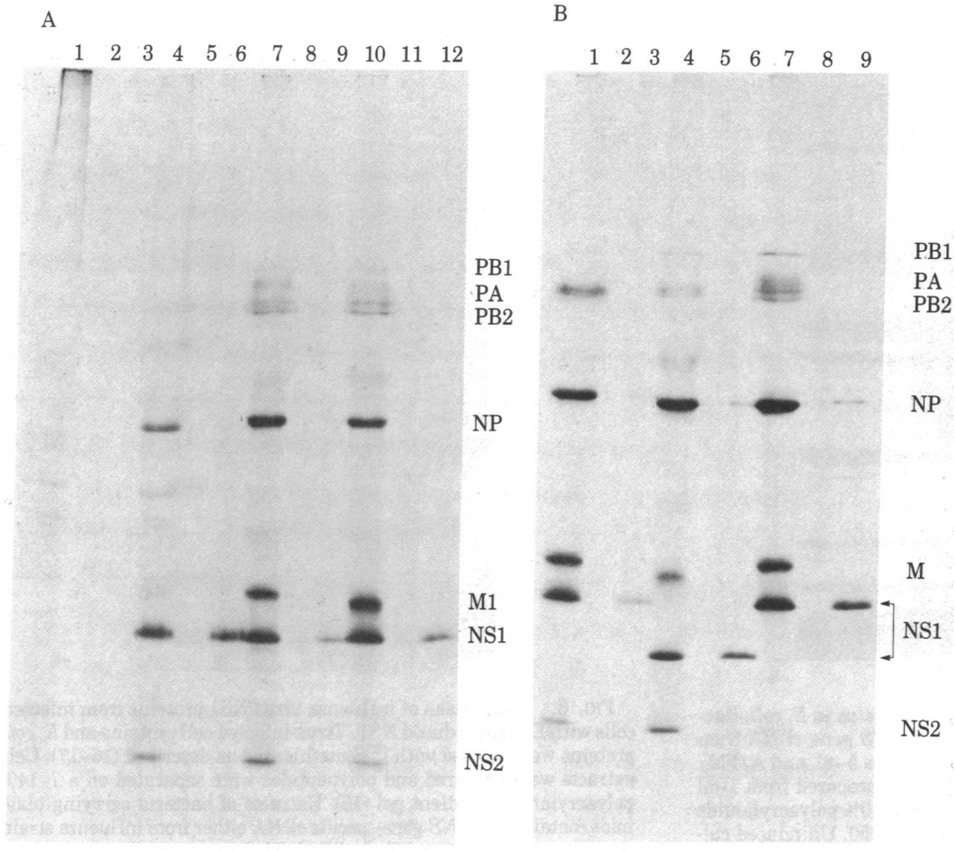

FIG. 4. Immunoprecipitation of influenza virus NS1 proteins using antiserum against *E. coli*-synthesized NS1 protein. Virus-infected MDCK cells labeled with [^{35}S]methionine were lysed and immunoprecipitated as described (23). (*A*) Mock-infected cells (lanes 1–3), cells infected with A/PR/8/34 virus (lanes 4–6), cells infected with A/FW/1/50 virus (lanes 7–9), and cells infected with A/USSR/90/77 virus (lanes 10–12). Total cell extracts (lanes 1, 4, 7, and 10), cell extracts reacted with normal nonimmune rabbit serum (lanes 2, 5, 8, and 11), and cell extracts reacted with anti-NS1 antiserum (lanes 3, 6, 9, and 12). (*B*) Cells infected with A/Udorn/72 virus (lanes 1–3), cells infected with A/FM/1/47 virus (lanes 4–6), and cells infected with A/duck/Alberta/60/76 virus (lanes 7–9). Total cell extracts (lanes 1, 4, and 7), cell extracts reacted with normal nonimmune rabbit serum (lanes 2, 5, and 8), and cell extracts reacted with anti-NS1 antiserum (lanes 3, 6, and 9).

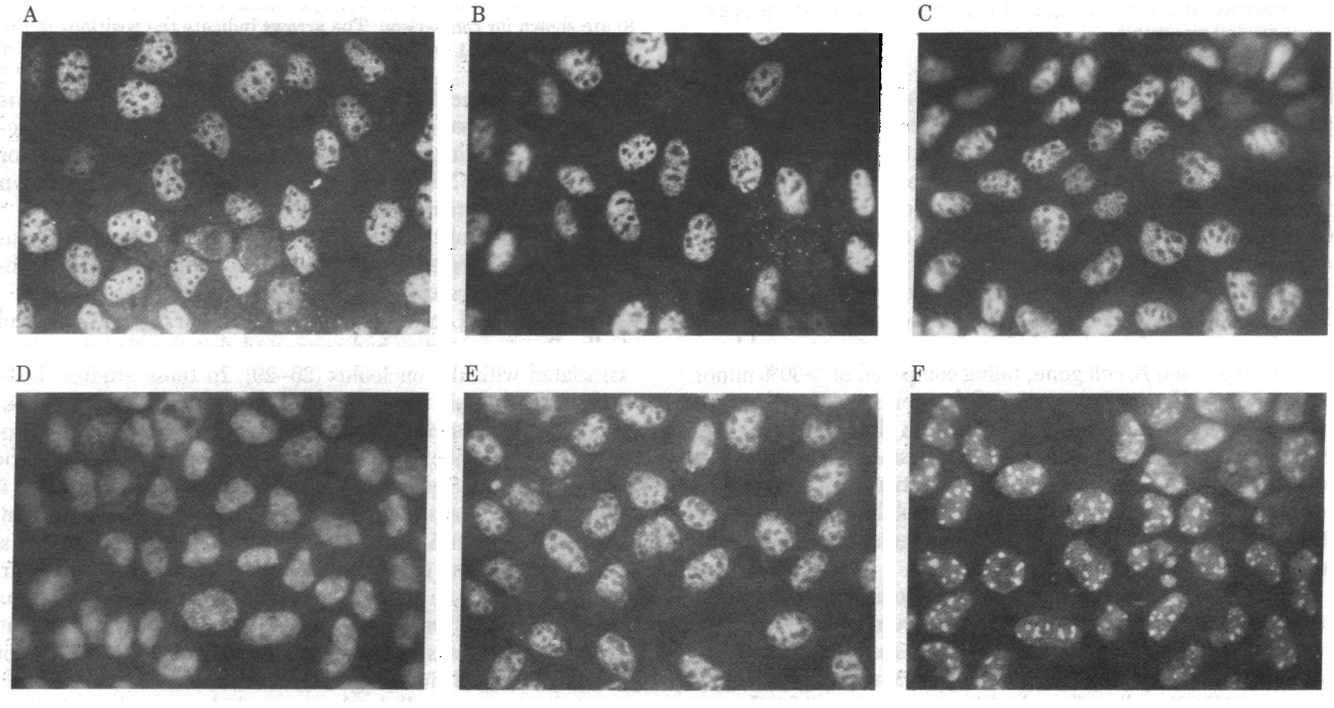

FIG. 5. Immunofluorescent staining of influenza virus-infected cells using NS1-specific antiserum. MDCK cells were grown on 13-mm glass coverslips and infected with influenza virus. Cells were harvested at the appropriate times, rinsed with phosphate-buffered saline, fixed in cold acetone (−20°C) for 10 min, and then air-dried at room temperature. Coverslips were treated with rabbit anti-NS1 antiserum (1/100 in phosphate-buffered saline) for 30 min at 37°C. Cells were then washed three times with phosphate-buffered saline and stained with fluorescein-conjugated goat anti-rabbit 7S gamma globulin (Hyland Diagnostics, Dearfield, IL) for 30 min at 37°C. Coverslips were then washed three times with phosphate-buffered saline and mounted on microscope slides with phosphate-buffered glycerol. Cells were examined and photographed using a Leitz fluorescence microscope. No staining was observed in uninfected cells (data not shown). Cells were infected with A/PR/8/34 virus and harvested at 2 hr (*A*), 4 hr (*C*), and 6 hr (*E*). Cells were infected with A/Bel/42 virus and harvested at 2 hr (*B*), 4 hr (*D*), and 6 hr (*F*). (×300.)

ences observed for the intracellular localization of the NS1 protein. In addition, because the maximum synthesis of NS1 protein in MDCK cells infected with either A/PR/8/34 virus or A/Bel/42 virus is during the first 3 hr of virus replication (unpublished observations), and all strains examined show nucleoplasmic staining at this time, it would appear that the primary site of NS1 activity is in the nucleoplasm.

CONCLUSION

This paper describes an *E. coli* expression system in which influenza virus proteins can be produced from cloned cDNA copies of the genomic RNA segments. An important feature of the pAS1 vector is that it is inducible and allows high level synthesis of proteins that might be toxic to cells (10). In the case of the NS1 protein, it was possible to obtain it as 20–25% of total cellular protein. In addition, individual viral proteins can be produced and purified free of other viral and eukaryotic host components for structural, functional, and immunological studies. This will be particularly important in obtaining the minor virion proteins (PB1, PB2, and PA) and the two minor nonstructural proteins (NS2 and M2) to provide antigens for the generation of monospecific antisera, which are not currently available for *in vitro* and *in vivo* studies.

This work was supported in part by grants from the National Institutes of Health (AI 11823 and AI 18998). J.F.Y. is a recipient of a Sinsheimer Scholar Award and P.P. is a recipient of an I. T. Hirschl Career Research Award.

1. Palese, P. (1977) *Cell* **10**, 1–10.
2. Lamb, R. A. (1983) in *The Genetics of Influenza Viruses*, eds. Palese, P. & Kingsbury, D. W. (Springer, Vienna), pp. 21–69.
3. Palese, P. & Kingsbury, D. W. eds. (1983) *The Genetics of Influenza Viruses* (Springer, Vienna).
4. Rosenberg, M., Ho, Y.-S. & Shatzman, A. (1983) *Methods Enzymol.* **101**, 123–138.
5. Baez, M., Taussig, R., Zazra, J. J., Young, J. F., Palese, P., Reisfeld, A. & Skalka, A. (1980) *Nucleic Acids Res.* **8**, 5845–5858.
6. Krystal, M., Buonagurio, D., Young, J. F. & Palese, P. (1983) *J. Virol.* **45**, 547–554.
7. Maniatis, T., Fritsch, E. F. & Sambrook, J., eds. (1982) *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
8. Ho, Y.-S., Lewis, M. & Rosenberg, M. (1982) *J. Biol. Chem.* **257**, 9128–9134.
9. Shatzman, A., Ho, Y.-S. & Rosenberg, M. (1983) in *Experimental Manipulation of Gene Expression*, ed. Inouye, M. (Academic, New York), pp. 1–14.
10. Shimatake, H. & Rosenberg, M. (1981) *Nature (London)* **292**, 128–132.
11. Sussman, R. & Jacob, F. (1962) *C.R. Hebd. Seances Acad. Sci. Ser. A* **254**, 1517–1519.
12. Emtage, J. S., Tacon, W. C. A., Catlin, G. H., Jenkins, B., Porter, A. G. & Carey, N. H. (1980) *Nature (London)* **283**, 171–174.
13. Heiland, I. & Gething, M. J. (1981) *Nature (London)* **292**, 851–852.
14. Davis, A. R., Nayak, D. P., Ueda, M., Hiti, A. L., Dowbenko, D. & Kleid, D. G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5376–5380.
15. Davis, A. R., Bos, T., Ueda, M. & Nayak, D. P., Dowbenko, D. & Compans, R. W. (1983) *Gene* **21**, 273–284.
16. Young, J. F. & Palese, P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6547–6551.
17. Parvin, J., Young, J. F. & Palese, P. (1983) *Virology*, in press.
18. Ikemura, T. (1981) *J. Mol. Biol.* **146**, 1–21.
19. Ikemura, T. (1981) *J. Mol. Biol.* **151**, 389–409.
20. Grosjean, H. & Fiers, W. (1982) *Gene* **18**, 199–209.
21. Gouy, M. & Gautier, C. (1982) *Nucleic Acids Res.* **10**, 7055–7074.
22. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. (1982) *J. Mol. Biol.* **162**, 729–773.
23. Graves, P., Schulman, J. L., Young, J. F. & Palese, P. (1983) *Virology* **126**, 106–116.
24. Baez, M., Zazra, J. J., Elliot, R. M., Young, J. F. & Palese, P. (1981) *Virology* **113**, 397–402.
25. Lamb, R. A. & Lai, C. J. (1981) *Cell* **21**, 475–485.
26. Dimmock, N. J. (1969) *Virology* **39**, 224–234.
27. Shaw, M. W., Lamon, E. W. & Compans, R. W. (1981) *Infect. Immun.* **34**, 1065–1067.
28. Taylor, J. M., Hampson, A. W., Layton, J. E. & White, D. O. (1970) *Virology* **42**, 744–752.
29. Krug, R. M. & Soeiro, R. (1975) *Virology* **64**, 378–387.