Journal of Cellular Biochemistry 33:39–51 (1987)
Proteases in Biological Control and Biotechnology 121–133

Site-Specific Mutagenesis of cDNA Clones Expressing a Poliovirus Proteinase

Bert L. Semler, Victoria H. Johnson, Patricia Gillis Dewalt, and Mary Frances Ypma-Wong

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, California 92717

The cleavage of poliovirus precursor polypeptides occurs at specific amino acid pairs that are recognized by viral proteinases. Most of the polio-specific cleavages occur at glutamine-glycine (Q-G) pairs that are recognized by the viral-encoded proteinase 3C (formerly called P3-7c). In order to carry out a defined molecular genetic study of the enzymatic activity of protein 3C, we have made cDNA clones of the poliovirus genome. The cDNA region corresponding to protein 3C was inserted into an inducible bacterial expression vector. This recombinant plasmid (called pIN-III-C3-7c) utilizes the bacterial lipoprotein promoter to direct the synthesis of a precursor polypeptide that contains the amino acid sequence of protein 3C as well as the amino- and carboxy-terminal Q-G cleavage signals. These signals have been previously shown to allow autocatalytic production of protein 3C in bacteria transformed with plasmid pIN-III-C3-7c. We have taken advantage of the autocatalytic cleavage of 3C in a bacterial expression system to study the effects of site-specific mutagenesis on its proteolytic activity. One mutation that we have introduced into the cDNA region encoding 3C is a single amino acid insertion near the carboxy-terminal Q-G cleavage site. The mutant recombinant plasmid (designated pIN-III-C3-μ10) directs the synthesis of a bacterial-polio precursor polypeptide that is like the wild-type construct (pIN-III-C3-7c). However, unlike the wild-type precursor, the mutant precursor cannot undergo autocatalytic cleavage to generate the mature proteinase 3C. Rather, the precursor is able to carry out cleavage at the amino-terminal Q-G site but not at the carboxy-terminal site. Thus, we have generated an altered poliovirus proteinase that is still able to carry out at least part of its cleavage activities but is unable to be a suitable substrate for self-cleavage at its carboxy-terminal Q-G pair.

Key words: proteolytic processing, bacterial expression, fusion protein

The expression of poliovirus genetic information depends upon the activity of at least two viral-encoded proteinases that carry out specific cleavages of precursor polypeptides [1–4]. The precursor polypeptides of poliovirus are derived from a single translation unit encoded in the positive strand virion RNA. Such a translation

Received May 15, 1986; revised and accepted August 18, 1986.

product is expressed beginning at a 5′ initation codon (nucleotide 743) and extending for 2,209 consecutive amino acids encoded in the ~7.5-kb genome. Prior to termination of synthesis of a complete polypeptide (corresponding to the long open reading frame of polio RNA), two viral-encoded proteinases begin to make the specific cleavages that eventually result in the functional capsid and noncapsid polypeptides required for a productive infection. The regime of synthesis and processing of viral precursor polypeptides is followed by all of the Picornaviridae, of which poliovirus is the prototype and best studied example. The sites of proteolytic cleavage within the poliovirus precursor polyprotein have been identified by a combination of nucleotide sequencing and amino- and carboxy-terminal sequence analysis of cleaved polypeptides [3–8]. One of the viral proteinases, protein 3C, cleaves at glutamine-glycine (Q-G) sites in polio precursor polypeptides. The second viral proteinase, protein 2A, cleaves at tyrosine-glycine (Y-G) sites encoded in the viral polyprotein. However, the recognition signals for cleavage by 3C or 2A must extend beyond the amino acid pair being cleaved because there are four Q-G pairs and eight Y-G pairs that are apparently not used as cleavage sites within the polio polyprotein [9].

In order to determine the functional correlation between specific amino acid residues and the cleavage activity of viral proteinase 3C, we have initiated a molecular genetic analysis of cDNA clones that express a functional 3C protein in bacteria. The isolation of cDNA clones of poliovirus type 1 (Mahoney) genomic RNA has been described [10–12]. We have also described the isolation of a recombinant plasmid that directs the inducible expression of an *Escherichia coli* polio fusion polypeptide containing the protein 3C amino acid sequence [13]. This fusion polypeptide has the polio-specific Q-G cleavage activity that is responsible for the intramolecular cleavage (or the *trans* activity of one precursor on another) that results in the production of the authentic 3C polypeptide in bacteria. We have utilized the bacterial expression of the active poliovirus proteinase to assay the effects of site-directed mutations introduced into the 3C coding region. In this report we described a simple mutagenesis protocol that allows the insertion of a single amino acid residue into specific sites within the 3C polypeptide. We describe the altered cleavage activity of one such mutant viral proteinase and discuss the implications of how the altered protein can retain its proteolytic activity without being a completely functional substrate for its own enzymatic cleavage.

MATERIALS AND METHODS

Plasmids and Cloning Procedures

The origin of the bacterial expression plasmid pIN-III-C3-7c is given in references [13,14]. The origin of the small plasmid vector pMV7 is given in reference [18]. To construct pMV7-2.9, we isolated a 2.9-kb BamHI-EcoRI fragment from a plasmid containing an infectious cDNA copy of the poliovirus genome [12]. The BamHI site of this 3′ terminal fragment of polio was converted to an EcoRI site using synthetic linkers. The fragment was then cloned into the EcoRI site of pMV7 to yield the recombinant plasmid shown in Figure 1. The plasmids pHf10-13 and pHf10-20 are derivatives of pEV104, an infectious recombinant plasmid that produces more than 1,000 poliovirus plaque-forming units/μg of plasmid DNA after transfection into COS-1 cells [12]. The 617-bp BglII-AccI fragment from pMV7-2.9 μ10 was subcloned into the corresponding region (ie, between nucleotides 5601 and 6218) of plasmid

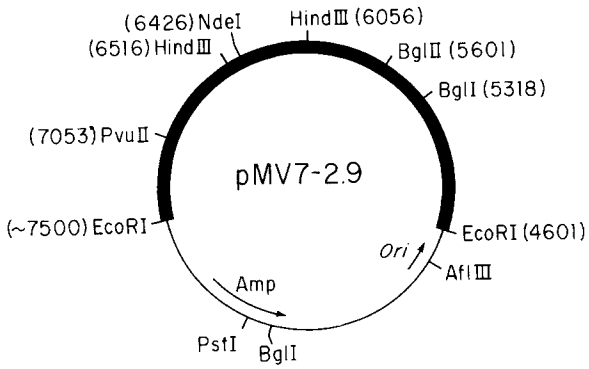

Fig. 1. Structure of polio recombinant plasmid pMV7-2.9. The numbers in parentheses refer to poliovirus type 1 (Mahoney) nucleotide numbers beginning from the 5′ end of the viral RNA. The details on the cloning and isolation of pMV7-2.9 are given in "Materials and Methods."

pEV104. Recombinant plasmids were screened for the absence of a Hinf I recognition sequence at nucleotide 5980 on the polio genome. The plasmids denoted pHf10-13 and pHf10-20 are derived from independent colony isolates of the same ligation and transformation steps. These two plasmids are identical in structure.

Generation and Screening of Hinf I Fill-in Mutations

To generate the Hinf I fill-in mutations, DNA from plasmid pMV7-2.9 was linearized with a random partial digest using restriction endonuclease Hinf I. Presumptive linear DNA fragments were isolated by preparative agarose gel electrophoresis and treated with the Klenow fragment of DNA polymerase I. The blunt-ended DNA fragment was then incubated with T4 DNA ligase and used to transform the C600 strain of *E coli*. Transformants were screened by rapid plasmid preparations [15] followed by digestion of isolated DNA with Hinf I. The DNA fragments generated by the above digestions were end-labeled with ^{32}P-α-dATP using the Klenow fragment of DNA polymerase I. Labeling reactions were carried out for 30 min at 25°C, and the labeled fragments were then separated by electrophoresis on an 8% polyacrylamide gel (40 cm × 20 cm × 1.5 mm).

Bacterial Expression Experiments

The induction and pulse-chase labeling of bacteria containing pIN-III-C3-7c recombinant plasmids were carried out as described by Hanecak et al [13,1]. Immunoprecipitation, sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis, and the origin of the polyclonal antiserum to protein 3C were as previously described [1,16].

DNA Transfections and Virus Plaque Assays

COS-1 cells were transfected with CsCl gradient-purified plasmid DNAs using the calcium phosphate precipitation method as previously described [17,18]. Plaque assays were carried out on 60-mm plates of HeLa cell monolayers under a semisolid agar overlay containing DME, 6% fetal calf serum, and 0.45% agarose. Cells incubated at 33.5°C were stained with crystal violet 3–4 days after infection, while those incubated at 37°C were stained 2–3 days after infection.

RESULTS

The strategy that we employed to obtain poliovirus proteinase mutants bearing a single amino acid insertion at a unique site is outlined in Table I. Briefly, we have carried out partial digestions with restriction endonuclease Hinf I on a plasmid containing a subgenomic cDNA insert of poliovirus RNA. The plasmid, called pMV7-2.9, contains a cDNA insert corresponding to nucleotides 4601 to 7500 (including the 3′ polyadenylate tract) on the polio genome. This region of the polio genome encodes part of the P2 region polypeptide and all of the P3 region polypeptides (refer to refs [3,4] for detailed polio processing maps). The insert has been cloned into the EcoRI site of a shortened derivative of plasmid pBR322, a 2.1-kb vector we have designated pMV7 [18]. A diagram of plasmid pMV7-2.9 is shown in Figure 1. After partial digestion of pMV7-2.9 DNA with Hinf I under conditions that produce DNA fragments containing 1 (or a few) randomly cleaved Hinf I site(s), the linear plasmid DNA was then isolated by preparative agarose gel electrophoresis and electroelution. The isolated full-length linear DNAs were then treated with the Klenow fragment of DNA polymerase I to fill in the three-nucleotide 5′ extension left by Hinf I cleavage, thereby yielding the insertion of a single amino acid residue at affected sites within the coding region of the polio polyprotein.

After ligation and transformation of *E coli,* individual bacterial transformants were screened for the loss of a single Hinf I recognition site in the pBR-polio chimeric plasmid DNAs. These plasmid DNAs were isolated by the rapid plasmid protocol of Birnboim and Doley [15] followed by digestion with restriction enconuclease Hinf I. The resulting DNA fragments were end-labeled in the presence of ^{32}P-α-dATP and DNA polymerase I (Klenow enzyme) and subjected to polyacrylamide gel electrophoresis. An example of the results that we obtained using the above analysis of the Hinf I mutagenesis transformants derived from pMV7-2.9 is shown in Figure 2. The electrophoretic profile of labeled Hinf I fragments from 12 randomly isolated transformants is displayed along with that of fragments generated from the parent plasmid (wt lane). From the sizes of fragments predicted by the nucleotide sequence of pMV7-2.9, it appears that the DNAs shown in lanes 3, 4, 5, 7, and 12 are derived from the parental plasmid without the loss of any of the Hinf I recognition sites. Such plasmid DNAs could be generated by the rejoining of linear, Hinf I-digested pMV7-2.9 DNAs that were not blunt-ended during the incubation with DNA polymerase I (Klenow) prior to the ligation and transformation steps.

TABLE I. Strategy for Isolation of Single Amino Acid Insertions in cDNA Clones of the Poliovirus Genome

1. Partial digestion of a subgenomic polio cDNA clone with restriction endonuclease Hinf I
2. Isolate linear plasmid DNA containing a single (random) Hinf I cleavage
3. Treat DNA with DNA polymerase I (Klenow) to fill in the three-nucleotide 5′ extension produced by Hinf I cleavage
4. Ligate and transform *E coli* with altered DNA
5. Screen for mutants using rapid plasmid preparations followed by Hinf I digestion and end-labeling
6. Analyze pattern of labeled DNA fragments by polyacrylamide gel electrophoresis
7. Correlate missing fragments and the presence of newly generated fragments to the known Hinf I digestion pattern of the wild-type plasmid
8. Subclone plasmid inserts containing single mutations into full-length polio cDNA constructs or bacterial expression vectors
9. Assay for expression of mutant phenotypes

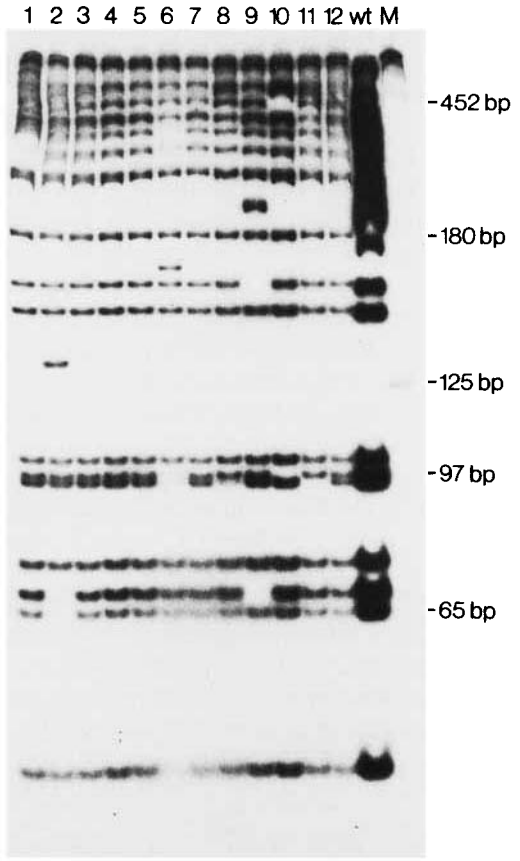

Fig. 2. Analysis of labeled DNA fragments from plasmids following the Hinf I mutagenesis protocol. The figure displays an autoradiograph of an 8% polyacrylamide gel after electrophoresis of ^{32}P-end-labeled restriction fragments. The numbered lanes refer to plasmid DNAs from different *E coli* transformants. The lane marked **wt** is unaltered plasmid DNA from pMV7-2.9 that was digested with Hinf I and end-labeled as described in "Materials and Methods." **Lane M** contains two end-labeled restriction fragments (564 bp and 125 bp) from a Hind III digest of bacteriophage lambda DNA.

In addition to the wild-type transformants described above, Figure 2 also shows the Hinf I restriction fragments from several altered derivatives of pMV7-2.9. For example, lane 2 has a restriction fragment pattern consistent with the loss of a Hinf I site in the region of the polio RNA polymerase (3D^{pol}), while lane 11 shows a pattern of Hinf I fragments that were produced from a plasmid lacking a Hinf I site in viral protein 2C (Semler and Johnson, unpublished data). Of particular interest to us is the pMV7-2.9 derivative that generated the Hinf I restriction fragment pattern shown in lane 10. This plasmid (designated pMV7-2.9-μ10) contains a filled-in Hinf I site at nucleotide 5980 in the polio genome that results in the loss of a 97-base-pair fragment and a 452-base-pair fragment and the generation of a new 549-base-pair fragment in the digest shown in Figure 2, lane 10. Because the site of mutagenesis (nucleotide 5980) in pMV7-2.9-μ10 was contained within the coding region for the viral proteinase 3C, we chose to examine the phenotypic expression of such a mutation in the experiments described below.

In order to study the effect of the mutation in pMV7-2.9-μ10 on protein processing, we subcloned a DNA fragment containing the filled-in Hinf I site at nucleotide 5980 into a bacterial expression vector. We have previously described [13] the cloning and expression of the poliovirus proteinase 3C in a plasmid called pIN-III-C3-7c (note that 3C is the new standardized designation for the viral proteinase known formerly as 7c [19]). The recombinant plasmid pIN-III-C3-7c is an inducible expression vector that uses the bacterial lipoprotein (*lpp*) promoter to direct the synthesis of an *lpp*-polio fusion polypeptide that contains the coding region for polypeptide 3C as well as the amino- and carboxy-terminal Q-G cleavage signals. These signals are used in *E coli* to allow autocatalytic cleavage of the precursor polypeptide into an authentic 3C polypeptide [13]. Thus, the bacterial expression system provides a very straightforward assay for the autocatalytic activity of viral proteinase 3C in cDNA clones that have been altered by site-specific mutagenesis.

The subcloning of an 816-bp HincII-HindIII fragment from pMV7-2.9-μ10 into the bacterial expression vector yielded a recombinant plasmid that we designated pIN-III-C3-μ10. The details of the cloning and the structure of the recombinant plasmid are given in reference [13]. The alteration in the 3C coding region that was introduced by the Hinf I mutagenesis protocol is shown in Figure 3. Note that the Hinf I digestion and fill-in with DNA polymerase I (Klenow) result in the insertion of a single serine residue at a site that is three amino acids upstream from the carboxy-terminal Q-G pair of polypeptide 3C. As a result, neither the poliovirus reading frame nor the Q-G cleavage site has been disturbed in the altered 3C construct.

The experiment that we carried out to determine the nature of the polio-specific polypeptides expressed in bacteria harboring the pIN-III-C3- μ10 plasmid was a pulse-chase labeling regime with ^{35}S-methionine followed by immunoprecipitation

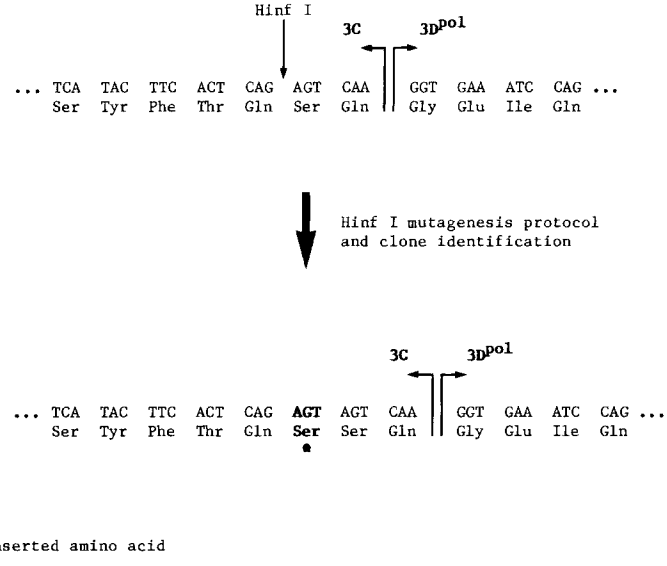

Fig. 3. Diagrammatic representation of the site of amino acid insertion in pIN-III-C3-μ10 DNA following the Hinf I mutagenesis protocol. The horizontal arrowheads refer to the carboxy-terminal coding region of viral proteinase 3C and the amino-terminal coding region of the viral RNA polymerase 3D^{pol}

with monospecific antiserum to protein 3C [13]. Exponentially growing *E coli* cultures were induced for 30 min with IPTG, pulsed with ^{35}S-methionine for 1 min, and then chased by the addition of excess unlabeled methionine for the times indicated in Figure 4. For the wild-type (WT) construct containing the unaltered 3C cDNA insert, the pattern of expressed polypeptides consists of 2 polypeptides that are cleaved products derived from the largest precursor (polypeptide 1). Polypeptide 1 is readily detectable only at the 1-min and 5-min time points and is chased into polypeptide 2 and polypeptide 3 at subsequent time points. By the 3-hr time point, there are approximately equivalent levels of polypeptides 2 and 3 in the induced bacteria. We have previously shown by immunoprecipitation and amino- and carboxy-terminal sequence analysis that both polypeptides 2 and 3 contain the authentic amino terminus of authentic polio protein 3C but differ in their carboxy-terminal amino acid residues [13]. As shown in Figure 5, polypeptide 3 corresponds to the authentic viral proteinase 3C, whereas polypeptide 2 has not been cleaved at the carboxy-terminal Q-G site. Note also that for the 1-min and 5-min time points shown in Figure 4, there is a

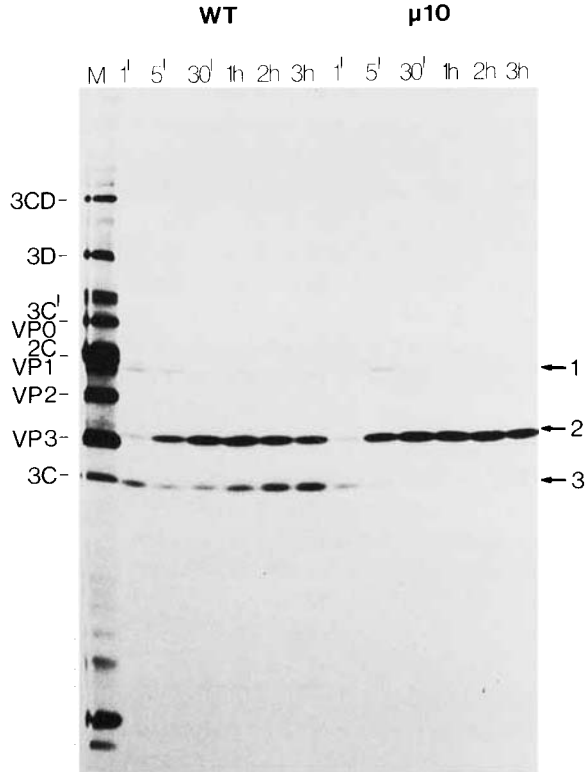

Fig. 4. Pulse-chase analysis of polypeptides produced in bacteria harboring the wild-type (**WT**) and mutant (**μ10**) 3C expression plasmids. Exponentially growing *E coli* were pulse-labeled with ^{35}S-methionine for 1 min and then chased with excess, unlabeled methionine for the times indicated. Following the chase, bacterial extracts were prepared and immunoprecipitated with α-3C serum. The immunoprecipitates were subjected to SDS-polyacrylamide gel electrophoresis on a 12.5% gel. The figure displays an autoradiograph of such a gel. **Lane M** contains an extract of poliovirus-infected HeLa cells that had been labeled with ^{35}S-methionine between 4 and 6 hr postinfection. The polio-specific proteins are indicated at the left of the figure. The numbers at the right of the figure refer to polio-specific polypeptides expressed in bacteria containing the recombinant plasmids.

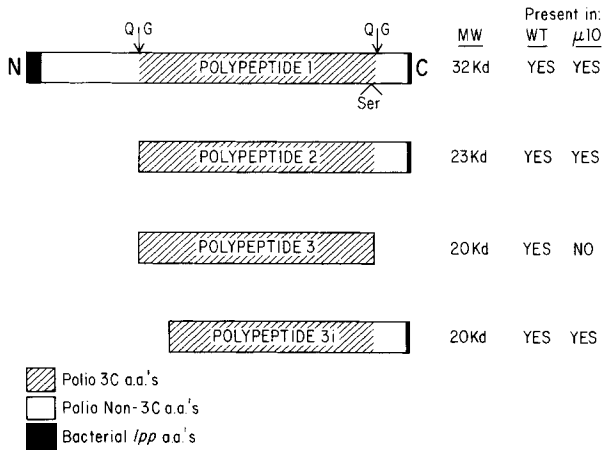

Fig. 5. Schematic representation of the polio-specific polypeptides expressed in *E coli* following induction of the pIN-III-C3-7c (**WT**) and μ10 plasmids. The serine insertion near the carboxy-terminal Q-G cleavage site of 3C is shown below polypeptide 1. The polio 3C amino acids are indicated with the cross-hatch lines. Polypeptide 3i is a product of an internal initiation within the 3C coding region and is not derived from any known polio-specific cleavage events.

polypeptide produced that is nearly identical in size to that of protein 3C but is generated too quickly in the pulse-chase experiment to be a cleaved product of the polypeptide 1 precursor. We have previously shown by amino-terminal sequence analysis that such a protein (called polypeptide 3i) is generated from an internal initiation site located 27 amino acid residues downstream from the amino-terminal glycine of protein 3C [13]. The methionine codon at the above position is preceded by a consensus Shine-Dalgarno [20] sequence thought to involve the binding of prokaryotic mRNAs to ribosomes. Polypeptide 3i is not cleaved at the carboxy-terminal Q-G of protein 3C (refer to Fig. 5) and is not stable for longer than 5 min during our pulse-chase experiment (Fig. 4), perhaps because it lacks the *lpp* signal sequence that may stabilize expressed proteins by inserting them into the outer bacterial membrane [13].

The results of an ^{35}S-methionine pulse-chase labeling experiment of proteins expressed by bacteria containing pIN-III-C3μ-10 are shown in the right half of Figure 4 (μ10). At the 1-min and 5-min time points, the pattern of polypeptides observed is similar to that seen in the WT samples, ie, polypeptides 1, 2, and 3i. However, unlike the proteins expressed by the WT plasmid, the μ10 polypeptides 1 and 2 are not ultimately processed to yield authentic 3C (polypeptide 3). Instead, when the unstable product of internal initiation (3i) disappears by the 30-min time point, there is no concomitant production of polypeptide 3. The polypeptide 1 precursor appears to be cleaved to yield polypeptide 2, but then no further cleavage events occur that would result in the production of polio protein 3C (polypeptide 3). As seen in Figure 4, even at the 3-hr time point, the only stable polio-specific protein detected is polypeptide 2.

As shown diagrammatically in Figure 5, the presence of the additional serine residue near the Q-G cleavage site at the carboxy terminus of 3C does not interfere with the production of polypeptide 2. Thus the cleavage activity contained within the 3C coding region is able to carry out the amino-terminal Q-G cleavage in μ10 with the same relative efficiency as that seen in WT (refer to Fig. 4). However, neither polypeptide 1 nor polypeptide 2 produced by the μ10 plasmid can undergo further

cleavage at the carboxy-terminal Q-G site that would lead to the production of polio protein 3C (polypeptide 3). Note that the production of polypeptide 3i is unaffected by the mutation in μ10 because this protein is not the product of any authentic polio-specific proteinase activity [13].

The Q-G cleavage site that is not recognized in bacteria expressing the pIN-III-C3-μ10 polypeptides is the cleavage site that is utilized during a normal poliovirus infection to generate the proteinase (3C) and the viral RNA polymerase ($3D^{pol}$). Because the 3C proteinase activity in the μ10 construct was not affected in making its amino-terminal cleavage, we assumed that the defect occurred as a result of altering the amino acid sequences surrounding the carboxy-terminal Q-G site (see "Discussion"). The polio cDNA in the pIN-III-C3-μ10 construct contained the coding sequences for only 25 amino acid residues downstream from the carboxy-terminal Q-G site of protein 3C. These 25 residues represent the amino-terminal amino acids of the $3D^{pol}$ protein. However, since the complete $3D^{pol}$ protein contains 461 amino acids [3], it is possible that the folding of a precursor containing only 25 residues downstream from the cleavage site may not present a completely authentic cleavage substrate for the 3C proteinase. In support of this possibility is our observation (shown in Fig 4 and ref [13]) that a significant fraction of the polypeptide 2 molecules synthesized by bacteria containing the WT plasmid construct is never cleaved at the carboxy-terminal Q-G site of 3C. To determine whether the presence of the full complement of authentic viral coding sequences would allow the altered 3C protein encoded in the μ10 plasmid to be cleaved at its carboxyl-terminal Q-G site, we subcloned the appropriate fragment from pMV7-2.9-μ10 into a full-length polio cDNA clone (see "Materials and Methods" for details). We have previously shown that a recombinant plasmid (pEV104) containing a full-length polio cDNA and eukaryotic gene transcription and replication signals could produce infectious poliovirus at relatively high levels [12]. The full-length clones that we isolated containing the μ10 serine insertion in 3C are called pHf10-13 and pHf10-20. These recombinant plasmids were isolated from different bacteria colonies from the same ligation and transformation and have identical structures. The purified plasmid DNAs were then used in transfection experiments to investigate whether the 3C mutation could lead to the production of infectious poliovirus. The results of direct plaque production assays following transfection are shown in Table II, part A. For these experiments, the mutant or wild-type (pEV104) plasmids were transfected into monolayers of COS and HeLa cells: Transfection of pEV104 into HeLa cells produced 50 plaque-forming units (pfu)/μg of DNA, while transfection into COS cells produced greater than 1,000 pfu/μg of pEV104 DNA. This latter increase in specific infectivity is due to the ability of COS cells to replicate and express pEV104 DNA to high levels, as previously described [12]. When transfections were carried out with pHf10-20, no virus plaques were observed in COS-1 or HeLa cells. Thus, within the limits of detection of the above assay, pHf10-20 is not infectious.

The transfection assay followed by agar overlay and subsequent detection of plaques is only sensitive enough to detect a mutant virus that might grow to a titer 1/1,000 of that obtained for wild-type poliovirus (refer to Table II). Any viable mutants that might produce virus yields that are more than 3 logs lower than those of wild-type virus would go undetected by such an assay. To increase the sensitivity of our assays for infectivity of plasmid DNAs after site-specific mutagenesis, we assayed the liquid overlay media (ie, without agar) harvested 4–6 days after transfection with

TABLE II. Poliovirus Plaque Production Following DNA Transfection*

Plasmid	Cell line	pfu/μg
A. Direct assay for plaques		
pHf10-20	COS-1	0
pEV104	COS-1	>1,000
pHf10-20	HeLa	0
pEV104	HeLa	50

Plasmid	Incubation temperature (°C)	Titer, $\log_{10}$ pfu/ml
B. Assay of supernatants of transfected cells		
pHf10-13	33.5	0
pHf10-20	37	0
pEV104	33.5	6.00
pEV104	37	6.45

*Cells were transfected with plasmid DNA and salmon sperm carrier DNA as previously described [17,18]. In B, assays were carried out on HeLa cell monolayers as described in "Materials and Methods" and in [18]. pfu, plaque-forming units.

polio cDNAs. The harvested media (undiluted and diluted) was analyzed for virus production by plaque assay on confluent monolayers of HeLa cells. We have previously employed such an approach to detect recombinant viruses generated from chimeric cDNAs that did not produce virus in the direct agar overlay assay [18]. The results of assays derived from pHf10-13, pHf10-20, and pEV104 transfections of COS cells are shown in Table II, part B. Note that no infectious virus could be detected for pHf10-13 or PHf10-20 and that the lack of virus production at 37°C could not be attributed to temperature-sensitive growth. By contrast, the wild-type plasmid (pEV104) produced approximately 6 logs of virus at both 33.5°C and 37°C. Given the increased sensitivity of this second assay method, we conclude that the 3C proteinase mutation contained in pHf10-13 and pHf10-20 leads to a lethal block in some aspect of the poliovirus replication cycle.

DISCUSSION

The experiments described in this communication provide evidence that the enzymatic activity of polio proteinase 3C can be separated from its ability to serve as a substrate for its own self-cleavage activity. The bacterial expression system that we have employed allows us to analyze the expression of site-specifically altered 3C proteins without the potential complications of pleiotropic effects that such mutations may cause following transfection of eukaryotic cells with full-length cDNAs. We have previously reported the precise amino- and carboxy-terminal sequences of the polio-related polypeptides expressed in bacteria containing the pIN-III-C3-7c plasmid. Knowledge of these sequences allows the relatively simple determination of the origin of various polypeptides expressed by mutant plasmids and hence the nature of the defect displayed by such mutants.

In the case of the mutation that we introduced into pIN-III-C3-μ10, we have been able to analyze the effect of an amino acid insertion on the ability of the 3C

activity to cleave at a site just downstream from the insertion. The μ10 plasmid directs the synthesis of a precursor protein (polypeptide 1) that can undergo cleavage at the amino-terminal Q-G site of polio protein 3C to produce polypeptide 2. Thus the serine insertion near the carboxy terminus of protein 3C has no effect on the self-cleavage of the viral proteinase to produce its authentic amino terminus. Implicit in this observation is the conclusion that such a mutation has not altered the ability of protein 3C to carry out Q-G cleavages.

The defect that results in no production of polypeptide 3 in bacteria harboring pIN-III-C3-μ10 is apparently caused by an altered cleavage signal at the carboxy-terminal cleavage site of polio protein 3C. Although the mutant precursor (polypeptide 1) still contains the authentic Q-G pair cleaved in the WT construct, the presence of an additional serine residue inserted three amino acids from this carboxy-terminal cleavage site produces pronounced biological effects. The inserted serine residue apparently disturbs part of the larger cleavage signal (other than Q-G pairs) that is recognized at that site. Alternatively, the presence of an additional amino acid may cause an altered secondary structure of the precursor polypeptides that makes the carboxy-terminal Q-G pair inaccessible to the active site of the enzyme. We may be able to differentiate between the above two possibilities by making similar amino acid insertions near other Q-G cleavage sites in the polio genome. Such changes should not uniformly produce the same effects if only an alteration of secondary structure is the cause of the noncleavage phenotype.

One consideration about protein-folding effects on cleavage site formation is whether the removal of polio protein 3C from its normal genomic location changes the contribution of the surrounding sequences in precursor polypeptides. These contributions may be important determinants in the folding of precursor polypeptides to generate substrates containing cleavable Q-G pairs. The experiments in which we analyzed the infectivity of full-length cDNAs containing the μ10 insertion (ie, pHf10-13 and pHf10-20) were designed, in part, to test the above possibility. Although the 3C activity expressed in bacteria harboring the pIN-III-C3-μ10 plasmid was unable to carry out the carboxy-terminal Q-G cleavage, it was possible that the removal of surrounding sequences (from viral protein $3D^{pol}$) coupled with the insertion of the serine residue (encoded at nucleotide 5980) was responsible for the lack of cleavage at this site. The plasmid constructs pHf10-13 and pHf10-20 were unable to produce infectious poliovirus following transfection into COS-1 cells. In addition, we could demonstrate that the block in virus production was not due to a temperature-sensitive phenotype since no virus was recovered at either 33.5°C or 37°C (refer to Table II). Because the 3C activity for Q-G sites has not been abolished in the proteins expressed by pIN-III-C3-μ10, we would anticipate that the activity would also not be affected following insertion of the altered DNA fragment from μ10 into full-length polio cDNA clones (ie, pHf10-13 and pHf10-20). Thus, the lack of infectivity of these latter two plasmids is probably not due to a general block in protein processing. Rather, the cleavage of the specific Q-G site that is affected by the μ10 insertion (at nucleotide 5980) must be an essential one for virus growth. The protein product (other than 3C) that is normally produced following cleavage at this Q-G pair is the RNA polymerase, $3D^{pol}$. The polymerase must apparently require the complete cleavage from precursor polypeptides in order for it to assume an active conformation capable of RNA polymerase activity. Such a conclusion is also supported by previous studies [21] that were unable to show detectable RNA polymerase activity in partially

purified extracts of protein 3CD (formerly called NCVP2 or P3-2). Protein 3CD is precisely the product that is obtained when the P3 region precursor is cleaved at the amino-terminal (but not the carboxy-terminal) Q-G site of protein 3C. It is also possible that the proteinase moiety of uncleaved (mutant) protein 3CD is not able to carry out other Q-G cleavages in a normal fashion, thereby resulting in no infectious virus production in cells transfected with pHf10-13 cDNA.

The approach to the mutagenesis of polio proteinase 3C that we have described in this report should provide a basis for the isolation of more in-phase mutations of this polypeptide. Such an approach has been recently employed to generate a viable poliovirus mutant in the 2A protein encoded in the middle of the viral genome [22]. It should be possible to use restriction enzymes other than Hinf I for three-nucleotide insertions in the 3C coding region. For example, both Dde I and Sau96 I leave three-nucleotide 5′ extensions, and both have two recognition sequences present in the 3C coding region. All of the above enzymes could also be used to generate sites for single amino acid deletions in the polio proteinase by removal of the three-nucleotide extensions using mung bean nuclease [23]. Such a varied array of genetically altered proteinase molecules will be useful in the ongoing analysis of the structure/function relationships of poliovirus protein processing.

ACKNOWLEDGMENTS

We are grateful to Dr. Ronnie Hanecak for helpful suggestions. This work was supported by grants from the American Cancer Society (MV-183) and the U.S. Public Health Service (AI 22693). P.G.D. and M.F.Y. are trainees of the U.S. Public Health Service (grants GM 07134 and CA 09054, respectively). B.L.S. is supported by a Faculty Research Award from the American Cancer Society.

REFERENCES

1. Hanecak R, Semler BL, Anderson CW, Wimmer E: Proc Natl Acad Sci USA 79:3973–3977, 1982.
2. Toyoda H, Nicklin MJH, Murray MG, Anderson CW, Dunn JJ, Studier FW, Wimmer E: Cell 45:761–770, 1986.
3. Kitamura N, Semler BL, Rothberg PG, Larsen GR, Adler CJ, Dorner AJ, Emini EA, Hanecak R, Lee JJ, van der Werf S, Anderson CW, Wimmer E: Nature 291:547–553, 1981.
4. Pallansch MA, Kew OM, Semler BL, Omilianowski DR, Anderson CW, Wimmer E, Rueckert RR: J Virol 49:873–880, 1984.
5. Semler BL, Anderson CW, Kitamura N, Rothberg PG, Wishart WL, Wimmer E: Proc Natl Acad Sci USA 78:3464–3468, 1981.
6. Semler BL, Hanecak R, Anderson CW, Wimmer E: Virology 114:589–594, 1981.
7. Larsen GR, Anderson CW, Dorner AJ, Semler BL, Wimmer E: J Virol 41:340–344, 1982.
8. Emini EA, Elzinga M, Wimmer E: J Virol 42:194–199, 1982.
9. Toyoda H, Nicklin MJH, Murray MG, Wimmer E: "Protein Engineering: Application in Science, Medicine and Industry." New York: Academic Press, Inc., in press.
10. Racaniello VR, Baltimore D: Proc Natl Acad Sci USA 78:4887–4891, 1981.
11. van der Werf S, Bregegere F, Kopecka H, Kitamura N, Rothberg PG, Kourilsky P, Wimmer E, Girard M: Proc Natl Acad Sci USA 78:5983–5987, 1981.
12. Semler BL, Dorner AJ, Wimmer E: Nucleic Acids Res 12:5123–5141, 1984.
13. Hanecak R, Semler BL, Ariga H, Anderson CW, Wimmer E: Cell 37:1063–1073, 1984.
14. Masui Y, Coleman J, Inouye M: "Experimental Manipulation of Gene Expression." New York: Academic Press, pp 15–32, 1983.
15. Birnboim HC, Doley J: Nucleic Acids Res 7:1513–1523, 1979.

16. Semler BL, Hanecak R, Dorner LF, Anderson CW, Wimmer E: Virology 126:624–635, 1983.
17. Graham FL, van der Eb AJ: Virology *52*:456–467, 1973.
18. Semler BL, Johnson VH, Tracy S: Proc Natl Acad Sci USA 83:1777–1781, 1986.
19. Rueckert RR, Wimmer E: J Virol 50:957–959, 1984.
20. Steitz JA, Jakes K: Proc Natl Acad Sci USA 72:4734–4738, 1975.
21. Van Dyke TA, Flanegan JB: J Virol 35:732–740, 1980.
22. Bernstein HD, Sonenberg N, Baltimore D: Mol Cell Biol 5:2913–2923, 1985.
23. Kowalski D, Kroeker WD, Laskowski M, Sr: Biochemistry 15:4457, 1976.